PHD

Three mathematical topics from the financial markets

Yousaf, Faisal A.

Award date:
2001

Awarding institution:
University of Bath

Three Mathematical Topics from the Financial Markets

submitted by

Faisal A. Yousaf

for the degree of Ph.D.

of the

University of Bath

2001

Signature of Author ....

Faisal A. Yousaf

Summary

This thesis is split into three distinct parts.

Part I investigates pricing a special type of barrier option which we have named a knock-down option. This is a discrete path dependent European option which is knocked out if the change in the daily closing log price of the underlying asset does not lie in some prescribed interval. Such an option does not appear to have a closed form solution. We present an elegant and remarkably accurate approximation for the price of this option appealing to results from classical probability theory. A put-call parity relationship is derived and a method for computing a numerically exact price of the option, using complex analysis and Fast Fourier Transform technology is also given. Numerical results are presented which provide a comparison of prices over different maturities, volatilities and intervals, investigating the accuracy of the approximations in each case.

Part II of this thesis is concerned with the potential approach to interest rate modelling advocated by Constantinides (1992) and Rogers (1997). The key element in this approach is to model the state-price density process. We begin by presenting a theoretical result which demonstrates how by choosing the underlying structure in the potential approach to be the Ornstein Uhlenbeck (OU) process, it is possible to replicate the bond prices given by Vasicek's spot rate model. Next we consider potential models which are driven by a finite state Markov chain. The focus here is inevitably numerical; we construct a framework for calibrating potential models to yield curve and exchange rate data. The objective is to develop a method to fit the term structure of several countries together with the exchange rates between them, the emphasis is placed on computationally feasible models. Empirical results are presented using sterling, US dollar and Deutschemark data.

Finally, in Part III of this thesis we consider a conventional equilibrium model and attempt to examine the nature of the changes which occur when a sudden unexpected exogenous shock is imposed to the market equilibrium. The market model considered has multiple agents who may invest in risky assets and a riskless deposit account and who each seek to maximise their expected integrated utility of consumption, which is taken to be the constant relative risk aversion or power-law utility function. The shock is imposed via the dynamics of the log Brownian dividend stream of one of the risky assets. We examine two examples and attempt to explain the nature of the changes which occur to the equilibrium portfolios of the agents and the endogenously determined interest rate as a result of the shock.

Acknowledgements

It's hard to know where to begin in acknowledging all the people who have helped me on the journey to completing this thesis.

- My supervisor Chris Rogers has been an immense help over the past three years. I would like to thank him for the patience he has shown as well as the advice and inspiration he has given me, I have enjoyed working with him immensely.

- I would like to thank my mum and gran for their love and support. They have encouraged and helped me through my years of study and I would not have been able to have done this without them.

The undergraduate years

- Dan Hobbs, Deborah Long and Jon Sharp are three good friends who started with me in the School of Mathematical Sciences at the University of Bath in October 1993. I had a great time doing my first degree with them.

- As an undergraduate I lived in several different places and made some wonderful friends in the process. I wish to mention the following people,
  - from my first year spent living in Eastwood House 9, Piers Baum, Grant Greenall, James Morris and Patrick Yeung;
  - from my second year spent living at 20 Seymour Road, Sarah Bailey and Neil MacDonald;
  - from my final year when I lived at 27 Holloway, Nic Brereton, Helen Grimes, Sarah Tilley and Claire Tyrrell.

- In 1995-96, I spent a year on an industrial placement at QuantiSci in Henley. I would like to thank Pete Grindrod, Dave Waters and Ado Metcalfe who were all at the company when I was there. I am grateful to Pete for encouraging me to undertake this PhD. Neil McAlpine was on placement with me and I have to thank him for keeping me sane over the year.

- Martin Oliver and Keith Walton have both helped me considerably and have offered valuable advice and support whenever I've needed it. I've enjoyed the many discussions we have had and I would like to thank them both.

The postgraduate years

Thank you to the EPSRC for funding my PhD. I also wish to acknowledge the following people:

- Ivana Beardmore, Jörg Berns-Müller, Graeme Boswell, Ciprian Coman, Nicoletta Del-Buono, David Esparza, Ceridwyn Fiddes, Dimitris Fouskakis, Stuart Hawkins, Vicky Henderson, Jon Heritage, Ramses Mena, Chris Sangwin, Robert Scheichl, Paul Spencer, Viveka Vivekanandan, Steven White, Andrew Whittle and Aaron Wilson, who were all postgrads in the department;
- Mary Baines, Jill Parker, Dena Pearson, Mark Willis and all the other members of administration and computer support;
- Véronique Brom, Bryce Grissom, Hamilton Jemmott, Altaf Moledina, Andy Selby, Giuseppe Spagnolo, Amy Wo and the Reds (Liverpool FC) who all helped take my mind off mathematics and make life more interesting;
- Mo Aydin, Claudia Bernardi and the rest of the Off-Campus Resident Tutor team, the Housekeeping staff in John Wood and the guys in the Security Office (I never enjoyed being woken up in the early hours of the morning because of fire alarms, parties and other problems but their humour and good nature made it bearable!);
- all the undergraduate students I have tutored over the years and in particular Simon Curzon, Emma Gosden, Alaistair Green, Marc Hussey, Laura Philips and Claire Turton who were part of a group I enjoyed tutoring for a whole year. I hope they enjoy their undergraduate degree at Bath as much as I did.

I would like to thank the other friends I have not listed who have helped make my time at the University of Bath so pleasant.

I thank God for giving me the opportunity to have done everything I have done.

April 2001

Contents

List of Figures

List of Tables

Chapter 1

Introduction to this Thesis

1.1 The World of Finance

The finance sector has developed and grown into one of the largest and most influential industries in the world today. Technological advances and the emergence of transnational corporations has ensured that whatever happens in one part of the world can be transmitted across the globe within seconds. Virtually the entire world economy has become a market economy and every nation is trying to utilise the opportunities this brings in order to help itself grow. As a result, the markets can now control the future of a large number of countries and so affect the day-to-day livelihood of many of their citizens.

One consequence of this is that there is an increased demand for mathematical and econometric tools which the corporate world and market participants can use to help gain an advantage in the industry.

This thesis considers three different problems from the world of finance and uses techniques from applied mathematics, probability theory, statistics and economics to develop solutions. The problems we look at originate from the financial markets and the emphasis throughout is to investigate practical solutions which may be of value in the real world.

1.2 Outline of this thesis

The thesis is split into three distinct parts. Part I contains just one chapter and focuses on pricing a special type of barrier option. We begin the chapter by presenting a brief introduction to basic option theory and the academic literature involving barrier options. Next we consider pricing a new type of exotic option which we term a knockdown option. This is a discrete path dependent European option which is knocked out if ever the change in the daily closing log price of the underlying asset does not lie in some prescribed interval. Such an option does not appear to have a closed form solution. We present an elegant and remarkably accurate approximation for the price of this option appealing to results from classical probability theory. A put-call parity relationship is derived and a method for computing a numerically exact price of the option, using complex analysis and Fast Fourier Transform technology is also given. Numerical results are presented and these provide a comparison of prices over different maturities, volatilities and intervals, investigating the accuracy of the approximations in each case. We conclude by examining the applicability and desirability of this option.

With the exception of the introductory and background sections, the whole of Chapter 2 of the thesis contains new material some of which is presented in the paper "Pricing a knockdown option", Yousaf [92].

Part II of this thesis is concerned with the potential approach to interest rate modelling. This is the largest of the three parts of the thesis and holds Chapters 3, 4, 5 and 6. Chapter 3 provides an introduction to interest rate modelling.

The focus in Chapter 4 is to examine the theory behind the potential approach. The majority of term structure models fall into one of two classes: those that model the spot rate process, or those that are concerned with the process of the forward rate. The potential approach is an alternative to these models which has received little attention from the mathematical finance community. This approach was advocated by Constantinides [34] and Rogers [84] and involves modelling the law of the state-price density process. In this chapter we focus on the theory given in Rogers [84] and present those concepts and results necessary to become familiar with this relatively new approach to interest rate modelling.

Chapters 5 and 6 contain new material. Chapter 5 begins by considering potential models which are driven by finite state Markov chains; this hasn't been considered before, and a discussion of the motivation behind this approach is provided. Explicit pricing expressions for zero coupon bonds, caps and swaps in this framework are then given. We also present a theoretical result which demonstrates how by choosing the underlying structure in the potential approach to be the Ornstein Uhlenbeck (OU) process, it is possible to replicate the bond prices given by Vasicek's spot rate model. Finally, a procedure for implementing this numerically using a discretised version of the OU process is given.

In Chapter 6 we construct a method for calibrating Markov chain potential models to yield curve and exchange rate data; the emphasis here is inevitably numerical. The objective of the work is not to provide an empirical comparison of this model against others, but to investigate the best method of calibrating and implementing Markov chain potential models. The framework we develop is capable of fitting the term structure of several countries together with the exchange rates between them. Empirical results are presented using yield curve and exchange rate data for sterling (GBP), the US dollar (USD), and the Deutschemark (DEM). Computational considerations are taken into account at all stages and the aim is to provide a method which is workable enough that it could feasibly be implemented by the industry.

Some of the material in Chapter 6 appears in the paper "Markov chains and the potential approach to modelling interest rates and exchange rates", Rogers & Yousaf [86].

Finally, Part III of this thesis contains Chapters 7 and 8. Here we consider a conventional equilibrium model and examine the changes which occur when a sudden exogenous shock is imposed to the market equilibrium. We start in Chapter 7 by giving the motivation for looking at this problem and then setting up the general multi-agent continuous time market model which we consider.

The economy we look at has one riskless deposit account and multiple risky assets. Each agent tries to maximise their expected net future utility, which is taken to be the power utility function. An explicit formula for the interest rate is derived in this chapter.

Chapter 8 considers two explicit examples where we impose exogenous shocks to the market via the dividend stream of the risky assets and analyse the effect this has on the endogenously determined interest rate and the portfolios of the various agents.

This work has also been presented in the paper "Examining the effects of shocks to a market equilibrium", Rogers & Yousaf [87].

1.3 A Guide to the Notation Used

The following abbreviations and notations are used in this thesis.

We use ':=' to mean 'is defined to equal' and '≡' denotes is equivalent to.

Abbreviations

- bp := basis point
- cdf := cumulative distribution function
- pdf := probability density function
- BM := Brownian Motion
- BOE := Bank Of England
- BRC := Base Rate Changes
- CI := Conditional Independence (calibration)
- CIR := Cox-Ingersoll-Ross (see §3.3.5)
- CRRA := Coefficient of Relative Risk Aversion
- DEM := Deutschemark
- EKDC := European KnockDown Call option
- EKDP := European KnockDown Put option
- ERM := Exchange Rate Mechanism
- FFT := Fast Fourier Transform
- GBP := British Pound (Sterling)
- HJM := Heath-Jarrow-Morton (see §3.4)
- IRD's := Interest Rate Derivatives
- MC := Markov Chain
- MLE := Maximum Likelihood Estimate
- NC := No Change
- OU := Ornstein Uhlenbeck
- RW := Random Walk (calibration)
- UK := United Kingdom

- USD := United States Dollar
- ZCB := Zero Coupon Bond

Notations

- $\mathbb{P}(X)$:= the probability of X
- $\mathbb{E}(X)$:= the expectation of X
- $\mathbb{V}(X)$:= the variance of X
- $\mathbb{N}$:= the natural numbers
- $\mathbb{R}$:= the real numbers
- $\mathbb{R}^+ := [0, \infty)$
- I_A := The indicator function of A
- $N(\mu, \sigma^2)$:= Normal distribution with mean μ and variance σ^2
- Φ := cdf of N(0,1) distribution
- $(a+b)^+ := \max(a+b, 0)$, for $a, b \in \mathbb{R}$.

Part I

Pricing A Knockdown Option

Chapter 2

Pricing a Knockdown Option

2.1 Barrier Options: theory and background

The simplest and most common type of option is a European call option.

Definition 2.1 European Call and Put Options
A **European call (respectively put) option** *written on some underlying asset provides the holder with the right, but not the obligation, to buy (respectively sell) a specified quantity of the underlying asset, for a prescribed amount (***exercise/strike price***), at some prescribed time (***exercise date***) in the future.*

It is important to appreciate that the holder of an European call (put) option has the right to buy (sell) the underlying, but they are not obliged to do so. If they choose not to, the option expires worthless. It is this fact that means that the option has some value, and a question which immediately arises is how to calculate its correct market value?

Consider a call option expiring at time T where the price process of the underlying asset at time t is $S(t) \equiv S_t$ and the exercise price is K. The payoff for this option at time T is

$$(S_T - K)^+.$$

In 1973 Black & Scholes [18] published a no-arbitrage pricing formula for such options using the assumption that the share price S_t follows

$$S_t = S_0 \exp\{\sigma W_t^* + (\mu - \frac{1}{2}\sigma^2)t\},$$

where W_t^* is a standard Brownian motion under the measure $\mathbb{P}^*$, S_0 is the share price at time-0, μ is the expected rate of return and σ is the volatility of the share.

The seminal papers of Harrison & Kreps [50] and Harrison & Pliska [51] show that the absence of arbitrage (suitably defined) is equivalent to the existence of a probability measure $\mathbb{P}$ (equivalent to $\mathbb{P}^*$) under which all discounted asset price processes are martingales (we will refer to this as a risk-neutral measure). Therefore, the time-t price of a European call option with expiry T, denoted by $C(t,T)$, is given by

$$C(t,T) = \mathbb{E}_t[e^{-r(T-t)}(S_T - K)^+] \tag{2.1}$$

where $\mathbb{E}_t \equiv \mathbb{E}[\cdot|\mathcal{F}_t]$ is the conditional expectation given the information $\mathcal{F}_t$ available at time-t taken in the measure $\mathbb{P}$, r is a constant representing the riskless interest rate. It can be shown that the explicit solution to (2.1) is given by the Black-Scholes formula (see for example Hull [4]),

$$C(t,T) = S_t\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2)$$

where

$$\begin{aligned} d_1 &= \frac{\log(S_t/K) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \\ d_2 &= \frac{\log(S_t/K) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}. \end{aligned}$$

Here $\Phi(.)$ is the cumulative normal distribution with zero mean and unit standard deviation. There are a number of excellent texts which discuss the issues of pricing derivatives, see for example Björk [1], Musiela & Rutkowski [7] and Neftci [8].

The field of option pricing has moved on considerably since the early work of Black, Scholes and Merton, options are now traded on just about every one of the world's major exchanges and as you might imagine, there are many types of options available. Classifications include American, Asian, Bermudan and many more, however, in this thesis we shall only consider European style options and so all references to options refer to this type.

As the theory has evolved, a large number of options with complicated contracts have been looked at in the literature; these are often referred to as exotic options. The work we shall do in this chapter will examine pricing a particular type of option from a class of options known as barrier options.

A large variety of barrrier options have been considered over the years, all concerned with whether or not the price of the underlying share crosses particular barriers. Barrier options are often used as hedging instruments and are appealing because they offer a cheap alternative to the plain European option. These options are very popular on the foreign exchange markets and extensive work has been carried out to obtain suitable pricing formulae for the many varieties. The simplest type of barrier option is the single constant barrier case. However, even here there are four cases to consider.

(a) Up-and-out single constant barrier option:
in this case the barrier is at a level $b > S_0$($\equiv$ initial share price) and if the price of the underlying asset crosses b at any time before expiry the option is worthless.

(b) Up-and-in single constant barrier option:
again the barrier is at $b > S_0$ but here the option is worthless unless the price of the underlying crosses b at some time before expiry.

(c) Down-and-out single constant barrier option:
this is similar to up-and-out except that the barrier is at $a < S_0$. The option is worthless if the price of the underlying crosses a at any time before the option's expiry.

(d) Down-and-in single constant barrier option:
here the barrier is at $a < S_0$. The option is worthless unless the price of the underlying crosses a at some time before expiry.

Merton [76] produced a pricing formula for a standard down-and-out European call option as far back as 1973 and the literature covers in some detail all of the above options. These options may sometimes be referred to as knock-out or knock-in options. Rubinstein & Reiner [88] give formulae for all the above types of single barrier options.

Double barrier knockout options also exist where the option is knocked out if the price of the underlying asset, S_t, goes above some upper barrier or below a lower barrier before the expiry of the option. Although several methods have been used to derive an analytical solution none have proved to be entirely satisfactory.

Closed form expressions for the prices of options are preferred, but when no such solution exists, it is often necessary to use approximation methods. Numerical methods for pricing barrier options are frequently used; Hull [4] and Boyle & Lau [20] may be consulted for further information. Standard lattice techniques are common and the binomial tree approach of Cox, Ross & Rubinstein (CRR) [38] is the most basic of the methods available. Binomial and trinomial trees can be used to price barrier options but as has been pointed out in a number of papers, the convergence can be slow and erratic leading to inaccurate prices, see for example Boyle & Lau [20]. The difficulty concerns the location of the lattice nodes near to the barrier. Rogers & Stapleton [81] provide an alternative binomial tree with random time steps which more accurately prices these types of options.

There are many articles in the literature on pricing options such as knock-outs, knock-ins and double barrier puts and calls. As well as those papers already mentioned, a small selection of further papers which the interested reader may wish to refer to include Broadie, Glasserman & Kou [26][27], Carr [31], Geman & Yor [47], Goldman, Sosin & Gatto [49], Kunitomo & Ikeda [68].

2.2 Introduction to the Knockdown Option

The remainder of this chapter is concerned with pricing an European put option with discrete path dependent knockout, which we name a knockdown option. We introduce an elegant and remarkably accurate approximation for the price of an option which is knocked out if ever the change in the daily closing log price of the underlying asset does not lie in some prescribed interval.

Much of the academic literature referred to in §2.1 relates to continuous barriers and hence assumes that there is continuous monitoring of the barrier. In reality this is rarely the case. Many real option contracts stipulate that only hourly prices, or daily closing prices may knock the option out, these clauses create what are in essence discontinuities in the barrier. Many of the existing models fail to address this problem and as a result tend to underprice the options. Discrepancies can be large, especially when a barrier is close to the starting price, see for example Kat & Verdonk [67] and Broadie, Glasserman & Kou [27]. The trade literature has for a long time voiced concern that these discontinuities are not fully taken into account by many of the models that are put forward, but it is only now, in relatively recent work, that this is being looked at in more detail.

The option that is considered in this chapter has a discrete path dependent knockout which operates if the change in the daily closing price exceeds specified upper and lower barriers. Such an option would appear to be a very desirable product, for both the investor and the market maker. As an example consider the situation where a trader has assumed a short position in a standard European put contract. With today's increasingly volatile economic climate this trader would be very susceptible to the large sudden drops in share prices that are becoming evermore frequent. The trader could stand to lose a lot of money! If however, a trader were to assume a short position in a European knockdown put option, they would have some significant security against these drops; a large sudden drop would just knock the option out and not mean large losses of money. Conversely, if we look at the situation from an investors point of view, there are also a number of perceived benefits. For example, investors no longer have to concern themselves with the freak intra-day fluctuations that would often knock out standard barrier options (especially if they were near the barrier). From this discussion, it follows that in liquid markets this knockdown option would essentially model the markets' volatility.

Unfortunately, one of the major obstacles in pricing discrete barrier options is that it is not possible to derive closed-form pricing results as in the fixed barrier continuous case. Moreover, the traditional way to proceed when pricing general barrier options has always been to adopt some form of tree method approximation, which, for a knockdown option, can be technically difficult to implement and also much slower than quasi-analytic solutions.

The approach we shall use here is to derive a (Black-Scholes-like) approximation for the price of our option. For this, we use a result from classical probability theory; the Central Limit Theorem. We then make use of a refinement to this theorem given in Petrov [9]; this forms the basis of §2.3. A note is also given at the end of that section stating and proving a put-call parity relationship for this option. §2.4 gives details of how to compute a numerically exact price for this option, utilizing results from complex analysis and fast Fourier transform technology. Numerical results are presented in §2.5 and we discuss the applicability of such an option and draw conclusions in the last two sections of this chapter.

2.3 Approximating the Price of a Knockdown Option

We assume that the price S of the underlying asset is a log Brownian motion and that the interest rate r is constant over the time period of the option, so:

$$X_t \equiv \log S_t \equiv \sigma W_t + (r - \sigma^2/2)t \equiv \sigma W_t + \mu t, \tag{2.2}$$

where W is a standard one-dimensional Brownian motion in the risk-neutral measure ($\mathbb{P}$) and σ is the volatility of S (also constant). Again the price of the asset at time t is denoted $S(t)$ and we may adopt the form S_t if no ambiguity arises.

To begin with, we restrict our attention to a European put option with expiry T and strike K, knocked out if ever the daily change in the log price is out of the closed interval $[a, b]$. We discretise the time interval dividing it into M equal pieces, so that $T = M\delta$, M being the number of days in the time period. We are interested in the end of day log prices; $X(j\delta) = \log S(j\delta)$, $j = 0, .., M$. Arbitrage pricing theory gives the time-0 price of this option as:

$$\mathbb{E}[e^{-rT}(K - S_T)^+;\ \zeta_j \in [a, b],\ j = 1, .., M], \tag{2.3}$$

where $\zeta_j = X(j\delta) - X((j-1)\delta)$, $j = 1, .., M$ and $\mathbb{E}[.]$ is the expectation taken in the risk-neutral measure.

2.3.1 The Central Limit Theorem Approximation

Equation (2.3) above can immediately be re-written as follows:

$$\text{EKDP} = \mathbb{E}[e^{-rT}(K - S_T)^+;\ \Gamma] = e^{-rT}\mathbb{E}[(K - S_T)^+ |\ \Gamma\ ]\mathbb{P}[\Gamma], \tag{2.4}$$

where $\Gamma \equiv (\zeta_j \in [a, b],\ j = 1, .., M)$[1].

The ζ_j are independent and $\zeta_j \sim N(\mu\delta, \sigma^2\delta)$ for $j = 1, .., M$, that is, for every j, ζ_j is normally distributed with mean $\mu\delta$ and variance $\sigma^2\delta$. Hence,

[1] EKDP := European knockdown put option

$$\begin{aligned}\mathbb{P}[\Gamma] &= (\mathbb{P}[\zeta_1 \in [a,b]])^M \\ &= \left[\frac{1}{\sigma\sqrt{2\pi\delta}}\int_a^b \exp(-\frac{(y-\mu\delta)^2}{2\sigma^2\delta})dy\right]^M \end{aligned} \tag{2.5}$$

Having obtained an explicit expression for one half of the right hand side of (2.4), we next consider $\mathbb{E}[(K - S_T)^+ | \, \Gamma]$. For this, we require the following lemma:

Lemma 2.1

(a) $\mathbb{E}[\zeta_1 | \, \zeta_1 \in [a,b]]$

$$= \left(\frac{1}{\mathbb{P}[\zeta_1 \in [a,b]]}\right)\left(\frac{\sigma\delta}{\sqrt{2\pi\delta}}\right)\left[\exp(-\frac{(a-\mu\delta)^2}{2\sigma^2\delta}) - \exp(-\frac{(b-\mu\delta)^2}{2\sigma^2\delta})\right] + \mu\delta.$$

(b) $\mathbb{E}[(\zeta_1)^2 | \, \zeta_1 \in [a,b]]$

$$\begin{aligned}&= \left(\frac{1}{\mathbb{P}[\zeta_1 \in [a,b]]}\right)\left(\frac{1}{\sigma\sqrt{2\pi\delta}}\right)\left[\sigma^2\delta a \exp(-\frac{(a-\mu\delta)^2}{2\sigma^2\delta}) - \sigma^2\delta b \exp(-\frac{(b-\mu\delta)^2}{2\sigma^2\delta})\right] \\ &\qquad + \sigma^2\delta + \mu\delta\mathbb{E}[\zeta_1 | \zeta_1 \in [a,b]].\end{aligned}$$

Proof

(a) Observe that,

$$\begin{aligned}&\mathbb{E}[\zeta_1 | \, \zeta_1 \in [a,b]] \\ &= \int_{\mathbb{R}} x\mathbb{P}[\zeta_1 \in dx | x \in [a,b]] \\ &= \int_{\mathbb{R}} x\frac{\mathbb{P}[\zeta_1 \in dx]}{\mathbb{P}[\zeta_1 \in [a,b]]} I_{\{x \in [a,b]\}} \\ &= \left(\frac{1}{\mathbb{P}[\zeta_1 \in [a,b]]}\right)\left(\frac{1}{\sigma\sqrt{2\pi\delta}}\right)\int_a^b x e^{\left(-\frac{(x-\mu\delta)^2}{2\sigma^2\delta}\right)} dx \\ &= \left(\frac{1}{\mathbb{P}[\zeta_1 \in [a,b]]}\right)\left(\frac{1}{\sigma\sqrt{2\pi\delta}}\right)\left[-\sigma^2\delta e^{\left(-\frac{(x-\mu\delta)^2}{2\sigma^2\delta}\right)}\right]_a^b + \mu\delta \\ &= \left(\frac{1}{\mathbb{P}[\zeta_1 \in [a,b]]}\right)\left(\frac{\sigma\delta}{\sqrt{2\pi\delta}}\right)\left[\exp(-\frac{(a-\mu\delta)^2}{2\sigma^2\delta}) - \exp(-\frac{(b-\mu\delta)^2}{2\sigma^2\delta})\right] + \mu\delta.\end{aligned}$$

(b) Analogously,

$$
\begin{aligned}
&\mathbb{E}[(\zeta_1)^2 |\ \zeta_1 \in [a,b]] \\
&= \int_{\mathbb{R}} x^2 \mathbb{P}[\zeta_1 \in dx | x \in [a,b]] \\
&= \int_{\mathbb{R}} x^2 \frac{\mathbb{P}[\zeta_1 \in dx]}{\mathbb{P}[\zeta_1 \in [a,b]]} I_{\{x \in [a,b]\}} \\
&= \left(\frac{1}{\mathbb{P}[\zeta_1 \in [a,b]]}\right)\left(\frac{1}{\sigma\sqrt{2\pi\delta}}\right) \int_a^b x^2 e^{\left(-\frac{(x-\mu\delta)^2}{2\sigma^2\delta}\right)} dx \\
&= \left(\frac{1}{\mathbb{P}[\zeta_1 \in [a,b]]}\right)\left(\frac{1}{\sigma\sqrt{2\pi\delta}}\right) \left[\int_a^b x(x-\mu\delta) e^{\left(-\frac{(x-\mu\delta)^2}{2\sigma^2\delta}\right)} dx + \int_a^b x\mu\delta e^{\left(-\frac{(x-\mu\delta)^2}{2\sigma^2\delta}\right)} dx\right] \\
&= \left(\frac{1}{\mathbb{P}[\zeta_1 \in [a,b]]}\right)\left(\frac{1}{\sigma\sqrt{2\pi\delta}}\right) \left[\int_a^b \sigma^2\delta e^{\left(-\frac{(x-\mu\delta)^2}{2\sigma^2\delta}\right)} dx - \left[\sigma^2\delta x e^{\left(-\frac{(x-\mu\delta)^2}{2\sigma^2\delta}\right)} dx\right]_a^b\right] \\
&\qquad + \mu\delta\mathbb{E}[\zeta_1 | \zeta_1 \in [a,b]] \\
&= \left(\frac{1}{\mathbb{P}[\zeta_1 \in [a,b]]}\right)\left(\frac{1}{\sigma\sqrt{2\pi\delta}}\right) \left[\sigma^2\delta a \exp(-\frac{(a-\mu\delta)^2}{2\sigma^2\delta}) - \sigma^2\delta b \exp(-\frac{(b-\mu\delta)^2}{2\sigma^2\delta})\right] \\
&\qquad + \sigma^2\delta + \mu\delta\mathbb{E}[\zeta_1 | \zeta_1 \in [a,b]].
\end{aligned}
$$

□

Next, using the Central Limit Theorem[2] we have that the law of $\sum_{j=1}^{M} \zeta_j$ given Γ is approximately $N(M\varepsilon, M\nu^2)$ where

$$
\begin{aligned}
\varepsilon &:= \mathbb{E}[\zeta_1 |\ \zeta_1 \in [a,b]], \\
\nu^2 &:= \mathbb{E}[(\zeta_1)^2 |\ \zeta_1 \in [a,b]] - \mathbb{E}[\zeta_1 |\ \zeta_1 \in [a,b]]^2
\end{aligned}
$$

which can be made explicit from lemma 2.1.

Hence

$$
\begin{aligned}
&\mathbb{E}[(K - S_T)^+ |\ \Gamma] \\
&\quad = \mathbb{E}[(K - S_0 e^{\sum_{j=1}^{M} \zeta_j})^+ |\ \Gamma] \\
&\quad \approx \left(\frac{1}{\nu\sqrt{2\pi M}}\right) \int_{-\infty}^{\log K/S_0} (K - S_0 e^w) \exp(-\frac{(w - M\varepsilon)^2}{2M\nu^2}) dw,
\end{aligned}
$$

[2] A statement of this theorem can be found in Williams [12] among many other places.

which after some standard algebra leads to

$$\mathbb{E}[(K - S_T)^+ | \ \Gamma] \approx K\Phi(d_1) - S_0 e^{\frac{1}{2}M\nu^2 + M\varepsilon}\Phi(d_2), \tag{2.6}$$

where, $\nu\sqrt{M}d_1 = (\log(K/S_0) - M\varepsilon)$, $\nu\sqrt{M}d_2 = (\log(K/S_0) - (M\nu^2 + M\varepsilon))$ and Φ is the standard normal cumulative distribution function.

Putting together (2.4), (2.5) and (2.6) gives us an initial approximation for the price of our option, namely;

$$\begin{aligned} \text{EKDP} &= \mathbb{E}[e^{-rT}(K - S_T)^+; \ \Gamma] \\ &\approx e^{-rT}\left[\frac{1}{\sigma\sqrt{2\pi\delta}}\int_a^b \exp(-\frac{(y-\mu\delta)^2}{2\sigma^2\delta})dy\right]^M [K\Phi(d_1) \\ &\qquad -S_0 e^{\frac{1}{2}M\nu^2 + M\varepsilon}\Phi(d_2)\Big]. \end{aligned} \tag{2.7}$$

Analogous arguments will allow us to derive a similar expression for the price of a European knockdown call option. It is also possible to derive a put-call parity relationship and we do this in §2.3.3.

2.3.2 Petrov's Refinement of the Central Limit Theorem

The above pricing expression clearly relies on the accuracy of the Central Limit Theorem approximation which we have used. An obvious question to address next is that of how to improve our approximation for the price. For this we turn to Petrov [9], Chapter 5 for an account of Osipov's asymptotic expansion of the Central Limit Theorem; we will make use the following refinement.

Theorem 2.1 (Theorem 5.22 Petrov [9]) *If we let $\{Y_n\}$ be a sequence of independent identically distributed random variables with mean $m(\Delta)$ and variance $\sigma(\Delta)^2$ then for $Z_n = Y_1 + Y_2 + .. + Y_n$ we have*

$$\mathbb{P}\left[\frac{Z_n - nm(\Delta)}{\sigma(\Delta)\sqrt{n}} \le x\right] = \Phi(x) + \frac{\alpha(1-x^2)e^{-\frac{x^2}{2}}}{\sqrt{72\pi n}} + o(n^{-\frac{1}{2}})$$

where

$$\alpha = \mathbb{E}\left[\left(\frac{Y_1 - m(\Delta)}{\sigma(\Delta)}\right)^3\right].$$

For a proof of this theorem the interested reader is referred to the original text of Petrov.

Using theorem 2.1 in our set up gives us the following refinement to our approximation:

$$\mathbb{P}(\sum_{j=1}^{M} \zeta_j^* \leq y) = \Phi(\frac{y - M\varepsilon}{\nu\sqrt{M}}) + \frac{\alpha}{\sqrt{72\pi M}}\left(1 - \frac{(y - M\varepsilon)^2}{\nu^2 M}\right) e^{(-\frac{(y-M\varepsilon)^2}{2M\nu^2})} + o(M^{-1/2}) \quad (2.8)$$

where

$$\alpha = \mathbb{E}\left[\left(\frac{\zeta_1 - \varepsilon}{\nu}\right)^3 \middle| \zeta_1 \in [a, b]\right],$$

and $\zeta_j^* = (\zeta_j | \zeta_j \in [a, b])$ for $j = 1, .., M$.

By calculating $\mathbb{E}[(\zeta_1)^3 | \ \zeta_1 \in [a, b]]$ and using the results of lemma 1, we see that

$$\begin{aligned} \alpha &= \left(\frac{\nu^{-3}}{\mathbb{P}[\zeta_1 \in [a, b]]}\right)\left(\frac{1}{\sigma\sqrt{2\pi\delta}}\right)\left[\sigma^2\delta a^2 \exp(-\frac{(a - \mu\delta)^2}{2\sigma^2\delta}) - \sigma^2\delta b^2 \exp(-\frac{(b - \mu\delta)^2}{2\sigma^2\delta})\right] \\ &\quad + \left(\frac{1}{\nu^3}\right)\left[2\varepsilon(\varepsilon^2 + \sigma^2\delta) + (\mu\delta - 3\varepsilon)\mathbb{E}[(\zeta_1)^2 | \zeta_1 \in [a, b]]\right]. \end{aligned} \quad (2.9)$$

Therefore, using (2.8) we have an extra term in our approximation (2.7); lengthy calculus gives the new term as:

$$\begin{aligned} &\left(\frac{\alpha(M\nu^2)}{\sqrt{72\pi M}}\right)\left[(1 + \frac{(\log(K/S_0) - M\varepsilon)}{M\nu^2})Ke^{-\frac{(\log(K/S_0)-M\varepsilon)^2}{2M\nu^2}}\right. \\ &\qquad \left. - S_0\sqrt{2\pi M\nu^2}e^{\frac{1}{2}M\nu^2 + M\varepsilon}\Phi(d_2)\right]. \end{aligned} \quad (2.10)$$

The expressions in (2.9) and (2.10) are verified in Appendix A.

So the approximate price for our option using Petrov's expansion of the Central Limit Theorem for independent identically distributed random variables is given by the sum of (2.7) and (2.10).

We ignore higher order terms in this expansion as they make no appreciable difference to the price of our option, as later numerical evidence will show. The reason for this is straightforward; the value of M that would be used in practice is likely to be high, options of less that a month (corresponding to $M = 21$) are rare. In fact, the theory seems to suggest that the refinement will be redundant for all options that have a large ($M > 40$) number of days before expiry.

2.3.3 Put-Call Parity Relationship

It is also possible to derive a put-call parity relationship for our option. This relationship will give an approximate price of a knockdown European call (denoted by c) in terms of the price of the corresponding approximate knockdown European put (denoted by p). Again, observe that arbitrage pricing gives the time-0 value of our European call option as:

$$\mathbb{E}[e^{-rT}(S_T - K)^+;\ \Gamma], \tag{2.11}$$

where

$$\Gamma \equiv (\zeta_j \in [a, b],\ j = 1, .., M).$$

Let us consider the difference in price between the put and the call on our option,

$$\begin{aligned}
p - c &= e^{-rT}\mathbb{E}[(K - S_T)^+;\ \Gamma] - e^{-rT}\mathbb{E}[(S_T - K)^+;\ \Gamma] \\
&= e^{-rT}\mathbb{E}[(K - S_T);\ \Gamma] \\
&= e^{-rT}K\mathbb{P}[\Gamma] - e^{-rT}\mathbb{E}[S_T;\ \Gamma] \\
&= e^{-rT}K\mathbb{P}[\Gamma] - e^{-rT}S_0\mathbb{E}[e^{\sum_{j=1}^{M}\zeta_j};\ \Gamma].
\end{aligned}$$

Then by the independence of the ζ_j, we deduce that,

$$
\begin{aligned}
p-c &= e^{-rT}K\mathbb{P}[\Gamma] - e^{-rT}S_0\left[\mathbb{E}[e^{\zeta_1};\ \zeta_1\in[a,b]]\right]^M \\
&= e^{-rT}K\mathbb{P}[\Gamma] - e^{-rT}S_0\left[\left(\frac{1}{\sigma\sqrt{2\pi\delta}}\right)\int_a^b e^x e^{-\frac{(x-\mu\delta)^2}{2\sigma^2\delta}}dx\right]^M, \\
&= e^{-rT}K\mathbb{P}[\zeta_1\in[a,b]]^M - e^{-rT}S_0\left[\left(\frac{1}{\sigma\sqrt{2\pi\delta}}\right)\int_a^b e^x e^{-\frac{(x-\mu\delta)^2}{2\sigma^2\delta}}dx\right]^M, \\
&= e^{-rT}K\mathbb{P}[\zeta_1\in[a,b]]^M - e^{-rT}S_0\left[\left(\frac{e^{(\frac{1}{2}\sigma^2\delta+\mu\delta)}}{\sigma\sqrt{2\pi\delta}}\right)\int_a^b e^{-\frac{(x-(\sigma^2\delta+\mu\delta))^2}{2\sigma^2\delta}}dx\right]^M,
\end{aligned}
$$

which then gives us the following put-call parity relationship:

$$
\begin{aligned}
p-c = e^{-rT}S_0\left[e^{\frac{M}{2}\sigma^2\delta+M\mu\delta}\right]\Bigg[\Phi\left(\frac{a-(\sigma^2\delta+\mu\delta)}{\sigma\sqrt{\delta}}\right) \\
-\Phi\left(\frac{b-(\sigma^2\delta+\mu\delta)}{\sigma\sqrt{\delta}}\right)\Bigg]^M + Ke^{-rT}\mathbb{P}[\zeta_1\in[a,b]]^M. \quad (2.12)
\end{aligned}
$$

2.4 Numerical Pricing of a Knockdown Option

Having found an expression for the approximate price of a knockdown option, our objective now is to establish how accurate this pricing formula actually is. As we do not believe that a closed form expression for the price of such an option exists, we are going to have to calculate an "exact" price via some kind of numerical method.

Consider pricing a discrete interval knockout put option.

$$
\begin{aligned}
\text{EKDP} &= \mathbb{E}[e^{-rT}(K-S_T)^+;\ \Gamma] \\
&= \mathbb{E}[e^{-rT}(K-S_0e^{\sum_{j=1}^M\zeta_j})^+;\ \zeta_j\in[a,b],\ \text{for } j=1,..,M] \\
&= \int_{-\infty}^{\infty} e^{-rT}(K-S_0e^x)^+g(x)dx \quad \text{where } g(x) \text{ is the density of } \left(\sum_{j=1}^M(\zeta_j;\ \zeta_j\in[a,b])\right) \\
&:= \int_{-\infty}^{\infty} e^{-rT}\psi(x)g(x)dx \qquad (2.13)
\end{aligned}
$$

We would now like to take Fourier transforms and so we need to use Parseval's identity on (2.13). However, since $\psi : x\mapsto (K-S_0e^x)^+$ is not in L^2, we cannot apply this identity directly, to proceed we instead consider the function $\psi_\varepsilon(x) = e^{\varepsilon x}\psi(x)$ for positive

ε. This then gives us that

$$\int_{-\infty}^{\infty} \{e^{\varepsilon x}(K - S_0 e^x)^+\} g(x) dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{\psi}_\varepsilon(\omega)\hat{g}(-\omega) d\omega,$$

where $\hat{\psi}_\varepsilon(\omega)$ is the Fourier transform of $\psi_\varepsilon(x)$,

$$\begin{aligned}\hat{\psi}_\varepsilon(\omega) &= \int_{-\infty}^{\infty} e^{ix\omega} e^{\varepsilon x}(K - S_0 e^x)^+ dx \\ &= \int_{-\infty}^{\log(K/S_0)} e^{x(\varepsilon + i\omega)}(K - S_0 e^x) dx \\ &= \frac{K e^{(\varepsilon+i\omega)\log(K/S_0)}}{(\varepsilon + i\omega)(\varepsilon + i\omega + 1)}\end{aligned}$$

and $\hat{g}(\omega)$ is the Fourier transform of $g(x)$,

$$\begin{aligned}\hat{g}(\omega) &= \int_{-\infty}^{\infty} g(x) e^{+i\omega x} dx \\ &= \mathbb{E}\left[\exp\left(+i\omega \sum_{j=1}^{M} \zeta_j\right); \zeta_j \in [a,b], \text{ for } j = 1,..,M\right] \\ &= \prod_{j=1}^{M} \left(\mathbb{E}[\exp(+i\omega\zeta_j); \zeta_j \in [a,b]]\right) \\ &= \left[\frac{1}{\sigma\sqrt{2\pi\delta}} \int_a^b e^{i\omega x} e^{\left(-\frac{(x-\mu\delta)^2}{2\sigma^2\delta}\right)} dx\right]^M.\end{aligned}$$

Hence

$$\mathbb{E}[(K - S_T)^+;\ \Gamma] = \lim_{\varepsilon \downarrow 0} \frac{K}{2\pi} \int_{-\infty}^{\infty} \frac{\hat{g}(-\omega) e^{\log(K/S_0)(\varepsilon + i\omega)}}{(\varepsilon + i\omega)(\varepsilon + i\omega + 1)} d\omega.$$

The singularity of the integrand at 0 causes problems, to overcome this, we make the substitution $z = \varepsilon + i\omega$ to give:

$$\mathbb{E}[(K - S_T)^+;\ \Gamma] = \lim_{\varepsilon \downarrow 0} \frac{K}{2\pi} \int_{\varepsilon - i\infty}^{\varepsilon + i\infty} \frac{\hat{g}(-\frac{(z-\varepsilon)}{i}) e^{z\log(K/S_0)}}{z(z+1)} \frac{dz}{i}.$$

Since $z \mapsto \frac{e^{z\log(K/S_0)}}{z(z+1)}$ is holomorphic in the right half plane, we can use Cauchy's Theorem for contour integrals[3], to integrate along any line in the right half plane. Picking $q \in (0, \infty)$, we have in the limit as $\varepsilon \downarrow 0$

$$\text{EKDP} = \mathbb{E}[e^{-rT}(K - S_T)^+;\ \Gamma] = \frac{Ke^{-rT}}{2\pi} \int_{q-i\infty}^{q+i\infty} \frac{\hat{g}(iz)e^{z\log(K/S_0)}}{z(z+1)} \frac{dz}{i},$$

the contribution from the distant ends of the rectangular contour tending to zero. This can then be evaluated using a numerical inverse Fourier transform routine, these are well documented, see for example Press, Teukolsy, Vettering & Flannery [10].

One outstanding question remains; how do we pick q? It turns out from numerical investigation that it matters little; convergence is rapid enough for a wide range of values of q. However, it is appropriate to note that there are sophisticated approaches to determine optimum values for q, the interested reader should refer to results on saddlepoint approximations (see for example Rogers & Zane [82], Daniels [39] and for a thorough treatment Jensen [5]).

2.5 Numerical Results

In this section, we will present prices for the knockdown option as computed by the Central Limit Theorem, the refinement to the Central Limit Theorem and using the numerical Fourier transform of §2.4. The price of the corresponding standard European put option with no knockout will also be given using the Black-Scholes formula (see §2.1).

The first observation we make is that the price of our exotic option will always be lower than that of the standard European option because of the possibility of it being knocked out. With this upper bound in mind we will compute prices for various options with differing intervals $[a, b]$, investigating the effect the interval, the volatility and the number of days to expiry has on the price of the option. In the results which follow, the interval is symmetric and centred at 0, therefore, an interval length $\zeta = 0.1$ implies the closed interval [-0.05,0.05].

2.5.1 Case A: a large number of days until expiry

The parameter settings which we use for this case are given in Table 2.1, notice that the length of the option (T) is 1 year and so we split the interval into 250 parts. This

[3]A full statement of this can be found in most books on complex analysis, see for example Marsden [6].

is probably the longest maturity knockdown option that we can have.

σ	r	S_0	K	T	M
0.25	0.06	100.0	120.0	1.0	250

Table 2.1: Parameter values for case A of numerical results for the knockdown option.

We consider the case where the permitted change in the daily log price is the same for each day over the life of the option, that is, ζ_j is the same for all j (we denote this value by ζ). We report how the value of ζ effects the price of a knockdown put option.

European Knockdown Put			
ζ	Computed Prices		
(Centre at 0)	CLT	Refinement	Exact
0.07	0.020137	0.020137	0.020137
0.08	1.035799	1.035788	1.035800
0.09	6.035960	6.035921	6.035968
0.10	12.383057	12.383013	12.383068
0.11	16.158328	16.158300	16.158336
0.12	17.668572	17.668558	17.668576
0.14	18.291919	18.291917	18.291921
0.16	18.333877	18.333877	18.333878
0.18	18.335755	18.335755	18.335756
0.20	18.335811	18.335811	18.335813
0.22	18.335813	18.335813	18.335814
Price of Option with No Knockout: 18.335814			

Table 2.2: Results for case A of knockdown options.

Examining the results in Table 2.2, we see that in this situation the Central Limit Theorem approximation achieves an accuracy of more than 1 part in 10^5, which is more than satisfactory for any trader.

It is also clear that the price of the knockdown option is very sensitive to the size of the end-of-day knockout interval. The plots in Figure 2-1 further illustrate how sensitive the price of our option is to the length of the interval ζ.

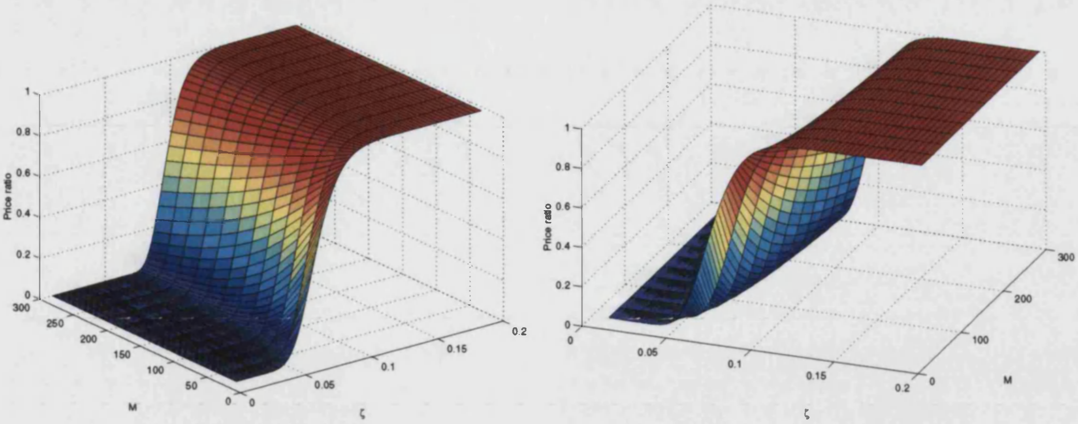

Figure 2-1: A comparison of the price of a European knockdown put option (computed by the Central Limit Theorem approximation) over the price of a European put with no knockout. We vary M, δ is fixed and $T = M\delta$. ζ is the length of the interval which is also varied. The parameter values taken are given in Table 2.1. We show the same surface plot from two different angles.

2.5.2 Case B: high volatility and a short time to expiry

In this case we will consider knockdown options which have a high volatility and a short maturity of only 25 days. The specific parameter settings which we shall use are given in Table 2.3.

σ	r	S_0	K	T	M
0.30	0.05	95.0	110.0	0.1	25

Table 2.3: Parameter values for case B of numerical results for the knockdown option.

European Knockdown Put			
ζ (Centre at 0)	Computed Prices		
	CLT	Refinement	Exact
0.07	2.734293	2.734288	2.734150
0.08	6.035462	6.035451	6.03500
0.09	9.418094	9.418078	9.417456
0.10	11.925163	11.925147	11.924587
0.11	13.413397	13.413384	13.413018
0.12	14.170225	14.170217	14.170058
0.14	14.656186	14.656183	14.656287
0.16	14.730206	14.730205	14.730394
0.18	14.738657	14.738657	14.738864
0.20	14.739390	14.739390	14.739599
0.22	14.739438	14.739438	14.739648
Price of Option with No Knockout: 14.739650			

Table 2.4: Results for case B of knockdown options.

In this example, (see Table 2.4) we find that the method which uses the refinement of the Central Limit Theorem is closer to the exact price than the prices generated using the standard Central Limit Theorem. This is in contrast to the results in case A and helps to emphasise the veracity of our statement in section (2.3.2) that the refinement is only of use if the number of days to expiry, M, is small.

2.5.3 Case C: how volatility effects the price

Next we will test our assertion that the price of a knockdown option would be sensitive to changes in the volatility. The parameter values we use are given in Table 2.5.

ζ	r	S_0	K	T	M
0.11	0.06	100.0	120.0	1.0	250

Table 2.5: Parameter values for case C of numerical results for the knockdown option.

We will look at the price for a fixed interval, ζ, centred at 0, for varying volatility. Observe from the results in Table 2.6 that for the interval ζ, as the volatility increases the price of the option plummets. This behaviour is expected, it is also interesting to compare it to the much more stable price of the corresponding vanilla put option (also given in Table 2.6). Figure 2-2 gives further diagnostics of this case but with varying maturities as well.

European put option				
σ	Knockdown prices			No Knockout
(volatility)	CLT	Refinement	Exact	B & S price
0.15	14.873704	14.873704	14.873706	14.873729
0.175	15.666059	15.666059	15.666061	15.668675
0.20	16.464717	16.464713	16.464717	16.521237
0.225	16.936597	16.936583	16.936596	17.414094
0.25	16.158328	16.158300	16.158336	18.335813
0.275	13.012962	13.012928	13.012994	19.278604
0.30	7.887758	7.887736	7.887807	20.236995
0.325	3.239001	3.238995	3.239038	21.207026
0.35	0.836529	0.836529	0.836544	22.185756

Table 2.6: Results for case C of knockdown options.

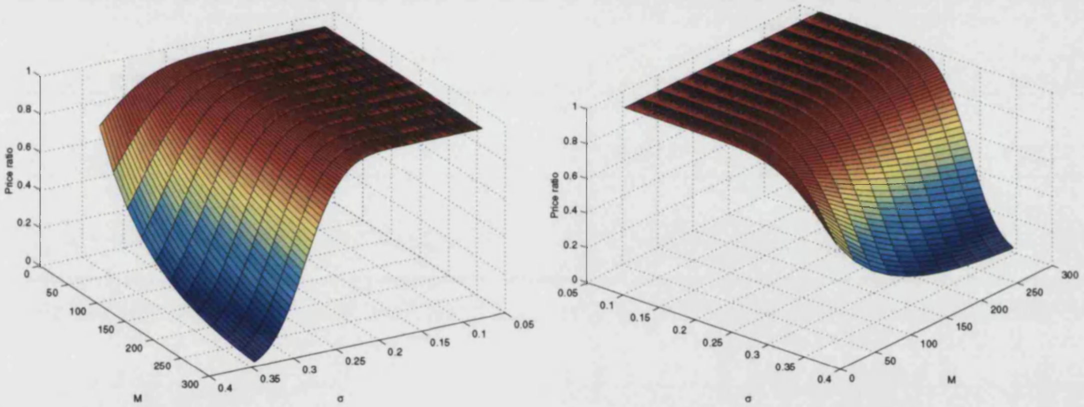

Figure 2-2: A comparison of the price of a European knockdown put option (computed by the Central Limit Theorem Approximation) over the price of a European put with no knockout, for varying M (δ is fixed and $T = M\delta$), and varying volatility, σ. The parameters used are given in Table 2.5. We give the surface plot from two angles.

2.5.4 Case D: usefulness of the approximation

Finally, one issue that is still unresolved is the usefulness of this approximation to price a knockdown option. In this case we shall investigate whether it is enough to simply estimate the price of the option as

$$\mathbb{P}[\zeta \in [a,b]]^M \mathbb{E}[(S_T - K)^+], \tag{2.14}$$

instead of our formula (2.7). To answer this question, we investigate the value of

$$\frac{\mathbb{E}[(S_T - K)^+;\ \Gamma]}{\mathbb{P}[\zeta \in [a,b]]^M \mathbb{E}[(S_T - K)^+]},$$

for different a, b and T. The results are shown in Figure 2-3, notice that the better the proposed estimate (2.14) to the price of a knockdown option, the closer the curve will be to a flat plane passing through $z = 1$ and parallel to the x-y axis.

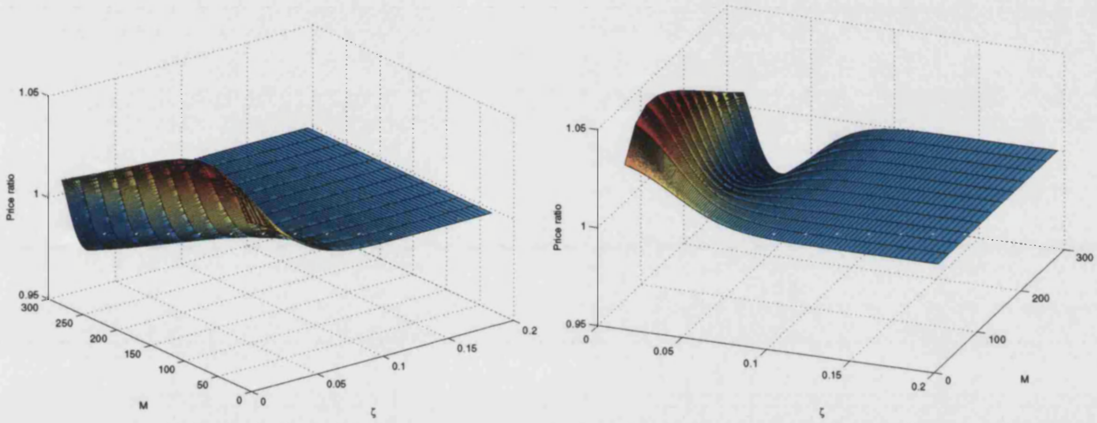

Figure 2-3: A plot depicting the ratio of our estimate (2.14) and the approximate price (2.7) for the knockdown put option with varying interval and days to expiry (fixed volatility). We show the same surface from two different angles.

2.6 Market Applicability

A recent review of leading academics and practitioners has suggested considerable interest in this type of option. The feedback we received indicated that such an option might have been traded a few years ago but we have been unable to confirm this and there appear to be no papers in the literature for this kind of option.

Nevertheless, it is clear that the characteristics of the knockdown option are appealing as they offer a degree of protection for both the trader and the investor. Such a product would probably be of particular interest to risk arbitrageurs who are naturally exposed to severe movements in the markets. Indeed, as we observed in §2.5.3 the price of this option is quite sensitive to volatility and so it could provide a way to bet on volatility if so desired.

Hedging a knockdown option is likely to be difficult; each day would behave as a near expiry, near the money digital option. Therefore delta and gamma hedging is not feasible as the positions taken would be huge. In practice the best option would be to hedge the option as if it was a standard no-knockout European option.

2.7 Conclusions

This chapter has been concerned with pricing a European put option with discrete path dependent knockout which we refer to as a knockdown option. One of the key features of the knockdown option is that the knockout condition is applied to the *daily change in the price* of the underlying share, and not, as is the case with most barrier options, dependent on the *price* of the underlying. This difference is crucial. Indeed, it is because of this that we have been able to price this option in a manner which is different to that normally adopted when pricing general barrier options.

Our study presents a straightforward expression for the approximate price of this option using the Central Limit Theorem as well as a refinement to this theorem given in Petrov [9]. We show how an exact price can be computed numerically using fast Fourier transform technology and this has allowed us to examine how accurate our approximation for the price actually is.

The results themselves indicate that the approximate price is accurate to the order of 1 part in 10^5. However, it should be recognised that this accuracy is dependent on the other parameters of the option and it can deteriorate to around 1 part in 10^4. In the absence of any closed-form exact pricing expressions for this option, this approximation is clearly an effective result.

The numerical results have also allowed us to verify the assertion made in our introduction that the price of the option would be sensitive to changes in market volatility.

Although the work in this chapter has been primarily concerned with pricing double sided knockdown options, this approximation may also be used to price a single sided knockdown option. In other words we could consider an option which is knocked out if the change in the end of day price is less than some specified level (or alternatively, greater than some specified level).

Part II

Interest Rate Modelling Using the Potential Approach

Chapter 3

Introduction and Brief Review of Interest Rate Modelling

The end of the twentieth century has seen a massive explosion in the variety and volume of interest rate related derivatives that are traded on the world's financial markets. In turn, this has lead to increased demands for mathematical models which can accurately price and hedge these new products.

As we start the 21st century, financial institutions, large corporations and major governments are still striving to develop complex computational and mathematical solutions to manage their exposure to interest rate risks and protect their assets in the event of market changes. However, even with the vast amount of research that has been carried out in this area, there are few clear cut solutions to the many problems they seek to solve. There is still little agreement on which model to use and considerable arguments over the relative merits of the different approaches.

The problems are so complex, the types of products traded so vast, and the world markets so complicated that no one model will ever be able to replicate market behaviour precisely. Nevertheless, academics and practitioners ceaselessly propose model after model, each with differing complexities and attributes and all attempting to please the practitioners.

The aim of this chapter is not to produce a complete survey of the literature over the past twenty years but more to give a brief introduction into the area of interest rate modelling and the terminology used within it. We shall concentrate our attention on one of the earliest interest rate models, the Vasicek model, and lay some of the foundations for the work that will be carried out in later chapters.

What is interest rate modelling?
The most obvious issue to address first is what is interest rate modelling and why study it? This question is probably best answered by posing and examining another question.

How much do I have to invest today to get \$10,000 in 5 years time?

This problem lies at the heart of any investigation into interest rates and the question is not only of interest to someone who is high up in the financial or corporate world! It is of relevance to anyone who has borrowed or invested money and that is most members of the general public.

Clearly, the exact details behind the solution to the above problem are not of direct interest to the general public. They know that if they wanted to guarantee a return of \$10,000 in 5 years time, they could simply go to the financial market and buy a bond maturing in 5 years for this amount. So the above question can be translated directly into the problem, for financial institutions, of pricing bonds.

Definition 3.1 Zero-Coupon Bonds[1]
A zero-coupon bond (ZCB) is a contract paying one unit of cash at a given date in the future, the **maturity** *date. A zero-coupon bond is also known as a discount bond and is termed 'zero-coupon' because there are no interim payments (coupons) during the life of the contract. Throughout we take ZCBs to be default free.*

To price such an instrument, we need to make assumptions about the evolution of interest rates over the period of the bond. In this case you cannot assume the interest rate is constant. It certainly isn't! The reason why we could make such an assumption in our work on pricing barrier options was because of the short lifespan of those options. However, the period under consideration for interest rate modelling can be anything up to 30 years.

In fact, nobody can reasonably forecast the future course of interest rates and so it is necessary to model them with random variables and therefore to develop stochastic models.

[1]There are slight variations in the definition of ZCB's in some texts; instead of one unit of cash they may be defined as paying some fixed, pre-specified, principle at maturity.

3.1 The Perfect Interest Rate Model

In developing a model for interest rates it helps to be clear about the characteristics we are looking for from our model. For this we use the list given in Rogers [83]: an interest rate model should be

- flexible enough to adapt to real-world influences;
- fast enough to be able to compute answers in real time;
- well-specified, so that required inputs can be observed or estimated;
- robust, in the sense that the model's results are insensitive to hard to measure parameters;
- realistic, in that the model will not do unwanted things e.g. negative interest rates;
- accurate, in that the model provides a good fit to data;
- tractable enough to provide an easy to handle regime for hedging interest rate derivatives.

There is no such thing as a perfect model and when designing a model it is often necessary to compromise between assumptions that are realistic and those that allow analytical tractability. One of the features which cannot be compromised is the need for fast responses; practitioners are not able to wait a long time for numerical routines to provide results, they need to respond to a host of different situations quickly. Therefore it is sometimes necessary to accept simple models due to computational restrictions, instead of more complicated and perhaps more realistic models.

The process followed by interest rates is considerably more complicated than that followed by a simple share price. An interest rate model must capture the behaviour of all the points on a yield curve, whereas, when modelling equity derivatives, you are only concerned with capturing the movements of a single point; the share price.

Another difference between interest rates and share prices is that interest rates appear to exhibit a behaviour known as **mean reversion.** That is, they have a drift which will tend to pull rates back to some long term average over time.

Therefore, when the interest rate is high, mean reversion tends to cause it to have a negative drift which will pull it down, and when the rate is low the opposite happens,

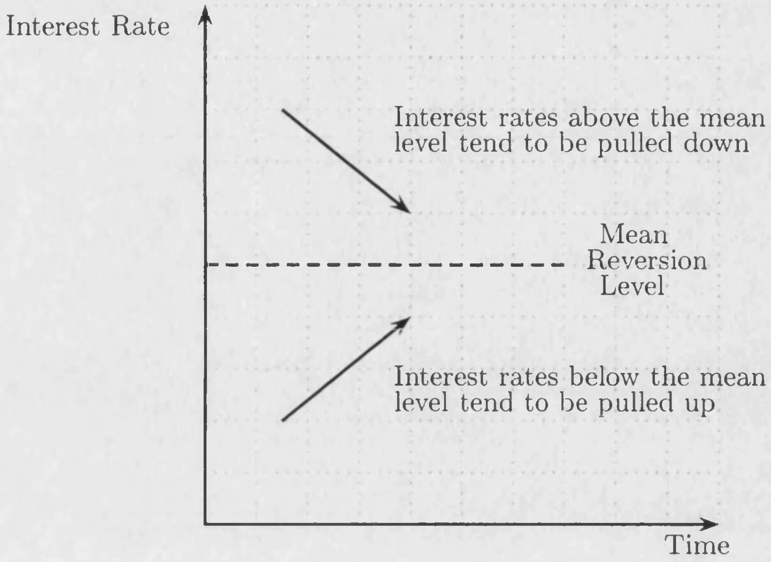

Figure 3-1: Mean Reversion

it has a positive drift pulling it up.

The economic arguments behind mean reversion are fairly straightforward: if interest rates are high, the economy tends to slow down and borrowers are less likely to want to borrow. The supply and demand argument then forces rates to decline. Conversely, when rates are low, borrowers are more inclined to borrow hence driving rates up.

3.2 The Yield Curve

We denote the time t price of a ZCB which matures at time T $(t < T)$ by $P(t,T)$.

Definition 3.2 Yield-to-Maturity
Let $Y(t,T)$ be defined by

$$Y(t,T) = \frac{-1}{T-t}\log P(t,T) \qquad \forall t \in [0,T). \tag{3.1}$$

$Y(t,T)$ is known as the yield-to-maturity of a ZCB maturing at time T.

Therefore the yield-to-maturity and the bond price are inversely related; when the bond price rises the yield falls and vice versa.

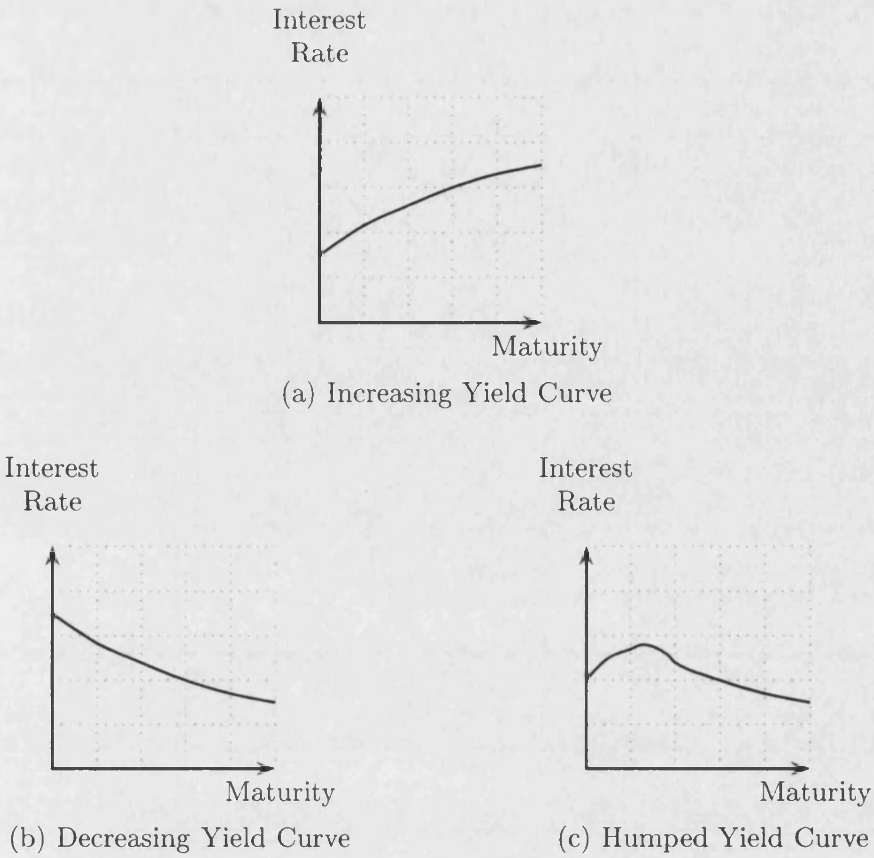

Figure 3-2: Classical Yield Curves

Definition 3.3 Yield Curve

The yield curve, also known as the term structure of interest rates, is the function that relates the yield $Y(t,T)$ to maturity T. Explicitly, the yield curve at time t is the function:

$$s \mapsto -\frac{1}{s}\log P(t,t+s) \qquad \forall t \in [0,T). \tag{3.2}$$

This theoretical definition cannot be applied directly at a practical level because of ZCB's and so in order to derive the yield curve in practice, it is necessary to take a number of liquid interest rate derivatives with different maturities. In this way, the yield curve on any given day is determined by the market and the prices quoted for that day.

Although yield curves have been known to come in a variety of different shapes, market data suggests that there are essentially three classical types of yield curve; increasing, decreasing and humped. These are illustrated in Figure 3-2.

Economists associate these shapes with different economic conditions.

1. An **increasing** curve is the most common of the three shapes and reflects an environment in which it is more rewarding to tie up money for a long time rather than a short time because long-term interest rates are higher than the short term rate. It is claimed that an economy with an increasing yield curve is one where the short rate is expected to rise.

2. Conversely, a **decreasing** or **humped** curve is often seen when the short rate is high but expected to fall.

For a more detailed analysis of yield curves the reader may consult Douglas [2].

3.3 Short-term Interest Rate Models

A substantial number of interest rate models proposed over the last 20 years have involved modelling the interest rate process directly.

Definition 3.4 Short-term interest rate
The short-term interest rate, also known as the spot interest rate, is the rate received on the shortest possible deposit over the next infinitesimal time period.

In practice, the spot rate is a very loosely defined quantity used to represent the instantaneous rate of riskless return at any time (i.e. the yield on a bond of infinitesimal maturity). It is often taken to be the three month LIBOR.

The spot rate approach to interest rate modelling still remains one of the most popular approaches used to hedge and price interest rate derivatives and there is a vast amount of literature on the subject.

Let us denote the short-term interest rate process by $(r_t)_{t \geq 0}$. Therefore, \$1 invested at time-t, which is continuously compounded will have grown to

$$\$ \exp\left(\int_t^T r_u du\right) \tag{3.3}$$

by time T. This is most commonly modelled as a continuous time process with continuous paths, though some models have allowed for jumps.

3.3.1 Pricing Zero-Coupon Bonds

If we denote the price at time $t \leq T$ of a ZCB of maturity T to be $P(t,T)$, then clearly $P(T,T) = 1$ and in the absence of coupon payments, $P(t,T) < P(T,T)$ $\forall t < T$. The last fact follows because people can invest in a risk free, non-negative, savings account, or carry cash at no cost.

The seminal papers of Harrison & Kreps [50] and Harrison & Pliska [51] show that the absence of arbitrage (suitably defined), is equivalent to the existence of a probability measure, $\mathbb{P}$, under which all discounted asset price processes are martingales. This measure is known as a martingale or risk-neutral measure.

Therefore, if we define

$$Z(t,T) := \exp(-\int_0^t r_u du)P(t,T) \tag{3.4}$$

$\forall t \in [0,T]$, then, under the martingale measure $\mathbb{P}$, we have that

$$Z(t,T) = \mathbb{E}[Z(T,T)|\mathcal{F}_t] \tag{3.5}$$

$\forall t \in [0,T]$. Hence,

$$P(t,T) = \mathbb{E}[\exp(-\int_t^T r_u)|\mathcal{F}_t] \tag{3.6}$$

$\forall t \in [0,T]$. This expression is the price of a ZCB as given by arbitrage pricing theory and it is frequently taken as the starting point for work in this area.

3.3.2 One-factor Models

Stochastic diffusion type models are often classified in terms of the number of factors in the model. This refers to the number of sources of uncertainty which is the dimensionality of the underlying Brownian Motion. In this section we consider two classical one-factor models of the short-term rate, the Vasicek Model [90] and the Cox-Ingersoll-Ross Model [37]. We will also briefly discuss the time-inhomogeneous extensions of these models proposed by Hull & White [58]. For each of these models, the underlying Brownian motion, W_t, is one dimensional, and this is the only source of uncertainty. Throughout what follows, the dynamics of r_t are specified under the risk-neutral measure $\mathbb{P}$.

3.3.3 The Vasicek Model

This is one of the earliest models for the short-term interest rate and was proposed by Oldrich Vasicek [90]. One of the appealing features of the Vasicek Model is its analytical tractability, the diffusion process is a mean reverting version of the Ornstein Uhlenbeck process;

$$dr_t = \beta(\rho - r_t)dt + \sigma dW_t \tag{3.7}$$

where ρ, β and σ are positive constants and W_t is a Brownian motion with respect to the risk-neutral measure. The analysis of this equation is straightforward.

(a) The mean of r_t: multiplying (3.7) by $\exp(\beta t)$ gives

$$e^{\beta t}dr_t \quad = \quad \sigma e^{\beta t}dW_t + \beta(\rho - r_t)e^{\beta t}dt$$

therefore,

$$d(r_t e^{\beta t}) = \sigma e^{\beta t}dW_t + \beta\rho e^{\beta t}dt.$$

In integral form this equation is

$$\begin{aligned} r_t &= r_0 e^{-\beta t} + e^{-\beta t}\int_0^t e^{\beta u}(\sigma dW_u + \rho\beta du) \\ &= r_0 e^{-\beta t} + \rho(1 - e^{-\beta t}) + e^{-\beta t}\int_0^t e^{\beta u}\sigma dW_u \end{aligned}$$

where r_0 is a positive constant. Taking expectations yields

$$\begin{aligned} \mathbb{E}[r_t] &= r_0 e^{-\beta t} + e^{-\beta t}\int_0^t \rho\beta e^{\beta u}du + \mathbb{E}[e^{-\beta t}\int_0^t \sigma e^{\beta u}dW_u] \\ &= e^{-\beta t}(r_0 + \rho e^{\beta t} - \rho). \end{aligned} \tag{3.8}$$

(b) The variance of r_t: for this part, we begin by applying Itô's formula to r_t^2

$$\begin{aligned} dr_t^2 &= 2r_t dr_t + d < r >_t \\ &= 2r_t\beta\rho dt - 2\beta r_t^2 dt + \sigma^2 dt + 2r_t\sigma dW_t, \end{aligned}$$

where $< r >_t$ is the quadratic variation of r_t. Taking expectations of the integral form of the above we get

$$\mathbb{E}[r_t^2] = r_0^2 + \int_0^t (2\rho\beta\mathbb{E}[r_u] + \sigma^2 - 2\beta\mathbb{E}[r_u^2])du.$$

Next, differentiating gives

$$\frac{d}{dt}\mathbb{E}[r_t^2] = 2\rho\beta\mathbb{E}[r_t] - 2\beta\mathbb{E}[r_t^2] + \sigma^2$$

which implies that

$$\frac{d}{dt}(e^{2\beta t}\mathbb{E}[r_t^2]) = 2\rho\beta e^{2\beta t}\mathbb{E}[r_t] + \sigma^2 e^{2\beta t}.$$

Integrating and using (3.8) yields

$$\mathbb{E}[r_t^2] = e^{-2\beta t}r_0^2 + \left[2\rho e^{-\beta t}(r_0 - \rho) + \rho^2 + \frac{\sigma^2}{2\beta} - \left(2\rho r_0 - \rho^2 + \frac{\sigma^2}{2\beta}\right)e^{-2\beta t}\right]. \tag{3.9}$$

Therefore (3.8) and (3.9) give

$$\mathrm{var}(r_t) = \frac{\sigma^2}{2\beta}\left(1 - e^{-2\beta t}\right). \tag{3.10}$$

(c) The limiting distribution of r_t: hence from the above we deduce that the limiting distribution of r_t is

$$N(\rho, \frac{\sigma^2}{2\beta}). \tag{3.11}$$

(d) Explicit formula for pricing bonds: we can derive an explicit expression for the bond price given by the Vasicek model. From the above calculations, with a little more calculus it quickly follows that,

$$\int_0^t r_u du \sim N(m_t, v_t^2) \tag{3.12}$$

where,

$$\begin{aligned} m_t &= \rho t + \frac{(1-\exp(-\beta t))}{\beta}(r_0 - \rho) \\ v_t^2 &= \frac{\sigma^2}{2\beta^3}[2\beta t - 3 + 4e^{-\beta t} - e^{-2\beta t}]. \end{aligned}$$

Then from (3.6) and (3.12) we get

$$\begin{aligned} P(0,t) \equiv \mathbb{E}[\exp(-\int_0^t r_u du)] &= \frac{1}{v_t\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(x-m_t)^2}{2v_t^2}} e^{-x} dx \\ &= \exp(-m_t + \frac{1}{2}v_t^2). \end{aligned} \tag{3.13}$$

(e) Remarks: the simplicity of the above expression is one of two very attractive features of Vasicek's model. The other one is its ability to provide a rich class of yield curves; this simple model is capable of producing all three of the classical yield curve shapes talked about earlier.

Although the Gaussian nature of Vasicek's model gives it an appealing tractability, it also leads to one of the more undesirable features of the model. That is, the possibility of returning negative interest rates. Negative rates clearly contradict our no-arbitrage condition and also contravene all empirical observations. It has often been argued that negative rates are not a big problem as the probability of these occurring is small. Rogers [83] demonstrates that this assumption is dangerous; we can choose sensible parameter values and still achieve negative interest rates!

(f) Undesirable example: by choosing ρ to be reasonably large when compared with the standard deviation $(\frac{\sigma}{\sqrt{2\beta}})$, we might imagine that we needn't worry about negative rates. However consider taking:

$$\rho = 0.1, \qquad \beta = 2 \times 10^{-3}, \qquad \sigma = 1.25 \times 10^{-3}.$$

Therefore we have:

1. $\rho > 5\frac{\sigma}{\sqrt{2\beta}}$

2. $\lim_{t\to\infty} \frac{1}{t} \log P(0,t) = 0.095$

but this implies $P(0,t) \to \exp(0.095t)$ which says that the bond price grows exponentially.

3.3.4 Hull and White's Extension of Vasicek's Model

A number of models with time-dependent parameters have been proposed over the years. The main reason for introducing time-dependency was because it allowed models to fit observed yield curves and volatilities perfectly. Moreover, models were also able to evolve so that they could more effectively reflect the cyclical nature of the economy and the impact of economic changes or changes in monetary policy.

In 1990, Hull & White [58] looked at a time inhomogeneous generalisation of Vasicek's model they considered the stochastic differential equation

$$dr_t = \sigma_t dW_t + \beta_t(\rho_t - r_t)dt \tag{3.14}$$

with ρ, β, $\sigma : \mathbb{R}^+ \to \mathbb{R}^+$ are locally bounded functions. We shall omit detailed analysis of this equation as it follows the same principle as that of Vasicek and it is not directly relevant to our later work. The interested reader should refer to the original Hull and White paper or alternatively Rogers [83] both of which derive explicit expressions of bond prices using this model.

Hull and White's extension of Vasicek's model can fit any yield-curve and any initial term-structure of volatility. However, there are impracticalities! If we estimated ρ, β and γ from today's yield curve data, then did the same in a week's time, the parameter values may well be completely different.

3.3.5 Cox-Ingersoll-Ross Model

The model proposed by Cox, Ingersoll & Ross [37], abbreviated to CIR, overcomes the problem of negative interest rates by considering a squared Gaussian model

$$dr_t = \beta(\rho - r_t)dt + \sigma\sqrt{r_t}dW_t \tag{3.15}$$

where ρ, β and σ are constants. This is another mean reverting stochastic differential equation with the added feature that if the interest rate starts non-negative then it remains non-negative (it could reach zero). Hence this model is more realistic.

A closed form expression for the price of a ZCB in the CIR framework can be derived in a similar fashion to that given for Vasicek's model in the previous section. However, as this will not be required for our later work the proof and explicit formulas are omitted from this text and the reader is referred to the original paper by Cox, Ingersoll and Ross [37]. Longstaff [69] also gives an expression for an option on a coupon bearing bond in the CIR framework.

Again, Hull and White have investigated a time inhomogeneous extension to CIR where the parameters ρ, $\beta, \sigma : \mathbb{R}^+ \to \mathbb{R}^+$ are time dependent locally bounded functions, see Hull & White [58].

3.3.6 Multifactor Models

One of the criticisms often levelled at one-factor Markovian models is that they imply perfect instantaneous correlation for the prices of bonds of different maturities. By this we mean that bond prices of different maturities all move in the same direction though not necessarily by the same amount. In practice the different maturities on the yield curve aren't perfectly correlated and this is perhaps one reason why multifactor models, which do not make this assumption, perform better in empirical studies.

The most common class of multifactor spot rate models are two factor models, these generally model the term structure by using the evolution of the short term interest rate together with some other economic variable to model the term structure. Brennan & Schwartz [24] [25], Duffie & Kan [42] and Longstaff & Schwartz [70] are just some of the papers that consider this class of model.

3.4 The Heath-Jarrow-Morton Class of Models

The Heath, Jarrow & Morton [52] [53] (HJM) approach to term structure modelling is based on modelling the motion of the forward interest rate curve. The HJM and spot rate paradigms are comfortably the most widely used approaches to interest rate modelling.

The basic idea in the HJM framework is to model the behaviour of the yield curve as a whole as opposed to modelling the short-term rate which is represented by a single point on this curve. These models are therefore sometimes referred to as whole yield curve models.

The HJM approach to interest rate modelling gives derivative prices which are determined by volatility parameters which describe the stochastic evolution of the entire yield curve.

3.5 Other Models

3.5.1 Discrete-time models

Although the models discussed so far have all been continuous-time there are a number of dominant discrete-time models that have helped to shape the direction of work in interest rate modelling.

The Ho & Lee [57] model leads the way. This binomial lattice model was one of the first to provide an exact fit to the current term structure of interest rates and is a discrete-time predecessor of the HJM model. Dybvig [44] and Jamshidian [62] both derived an equivalent continuous time formulation of this model:

$$dr_t = \mu(t)dt + \sigma dW_t.$$

Black, Derman & Toy (BDT) [17] also presented a binomial lattice model. Their model was not only able to fit the current term structure but also the current volatilities of interest rates. Jamshidian [63] derived a continuous time version for this model.

3.5.2 Market Models

Brace, Gatarek & Musiela [22] present a modelling methodology which directly uses market observables to build a model, these are often referred to as market models. This model is well suited to path-dependent derivative products and is relatively simple to calibrate.

3.6 Empirical Research

One of the major problems in interest rate modelling is that even with the large number of models that have been proposed, relatively little is known about how each of these models compare in terms of their ability to capture actual market behaviour. There is no common framework under which different models can be compared and tested to evaluate relative performance in a consistent way. The major obstacle is that the various modelling approaches are all very different, consequently, until recently very few papers have attempted to compare models and carry out empirical research to benchmark performances.

Chan, Karolyi, Longstaff & Sanders [32] begin to address this problem and compare a variety of continuous-time models of the short-term riskless rate using the Generalized Method of Moments. The results they obtain are interesting as their paper is the first to show evidence that the most successful models in capturing the dynamics of the short-term interest rate are those that allow the volatility of interest rate changes to be highly sensitive to the riskless rate.

There are other papers which test specific models. The first extensive empirical analysis of the CIR model was Brown & Dybvig [28]. They consider US Treasury bills, notes and bonds over the period 1952 to 1983. For each day they take the market price of the derivatives and estimate model parameters that produce the best least squares fit of model prices to observed market prices. This approach was also used by Brown & Schaefer [29], Chen & Scott [33] and Moraleda & Pelsser [78]. Gibbons & Ramaswamy [48] also conduct empirical tests of the CIR model of the term structure as do Heston [56], Longstaff [69], Longstaff & Schwartz [70] and Pearson & Sun [79].

Many interest rate models specify the stochastic behaviour of unobservable financial quantities such as instantaneous forward rates, the instantaneous short rate or its variance. This forces calibration procedures to transform these unobservable parameters to a set of market quantities and this can be problematic. One major criticism frequently made about interest rate models is that if a model is calibrated today, and then recalibrated even a day later, with only a small market change, the parameter values for the two calibrations will often be substantially different. If this happens it gives practitioners little faith in the modelling regime they are using.

The papers of Hull & White [59] [60] [61] consider the calibration of a certain class of short rate models to option prices. The models they consider are of the form

$$df(r_t) = \sigma(t)dW_t + [\theta(t) - a(t)f(r_t)]dt$$

where $f(r_t)$ is some function of the short rate r_t and W_t is a standard Brownian motion. The calibration of HJM type models are considered in Amin & Morton [14], Brace &

Musiela [21] and Brace, Garatek & Musiela [22].

Amin & Morton [14] test six term structure models in the HJM class. They use an implied volatility technique to estimate model parameters on a daily basis using Eurodollar futures and futures options data with different maturities over the period from 1987 to 1992.

The implied volatility technique is one that is frequently used by banks and other financial institutions to calibrate their HJM style interest rate models. The idea in this approach is to use current market prices of actively traded liquid derivatives to determine volatility parameter estimates and then to use these estimates to price the less liquid and more complicated exotic derivatives. One long held criticism of fits that use implied volatility is that invariably they involve a new estimation of the parameters on each trading day and this is often at odds with the theoretical foundations of the models which assume that the volatility parameters are constant over time. Nevertheless, this technique is commonly used and many academics and practitioners have accepted this apparent inconsistency.

Amin & Morton report results for pricing an average of 18 options each day using one and two parameter models. They quote impressive absolute errors for the prices of each option.

Finally, there have been two significant recent contributions to the literature on empirical comparisons of interest rate models. The first is a paper by Bühler, Uhrig-Homburg, Walter & Weber [30] which conducts a comprehensive comparison of forward and spot interest rate models. In this paper the authors consider 7 different models and use interest rate warrants on the German market from 1990-1993 to assess the performance of each model. They estimate the various model parameters using historical time series rather than using the implicit method described earlier. Bühler et al concede that it would not be feasible to use most of the models that they consider in a trading environment. This is primarily because their path-dependent nature make fast and efficient computation difficult.

The second recent paper on empirical comparisons is by Moraleda & Pelsser [78]. This paper also conducts comparisons between spot and forward interest rate models. However, in contrast to the work of Bühler et al, one of the primary objectives here is to consider models that are computationally tractable and so could be used by practitioners in a trading environment. Moraleda & Pelsser fit their chosen model to yield curve data and prices of caps and floors with different strikes and maturities. The data they use is from the US markets taken over the years 1993 and 1994. Their work uses an implied volatility technique to estimate parameters so that the sum of squares of differences between market and model prices is minimised. Some attention is also devoted to the performance of the models to out-of-sample data. For any given day they take parameter estimates of the previous day and compare the cap/floor prices produced by the models with the prices observed in the market. They only consider one-day out-of-sample tests.

The results in Moraleda & Pelsser's paper indicate that the spot rate models out perform forward rate ones both for in sample and out-of-sample data tests. This appears to be one of the best papers to-date on the calibration and fitting of interest rate models.

As a final note, it is worth remarking that purists often argue that the good fits obtained in empirical work involving implied volatility is primarily a result of the daily re-estimation of the volatility parameters and so it is not a good measure of the quality of the term structure models. Practitioners often respond to this criticism by commenting that their chief concern is to use models to price exotic derivatives and to revise daily hedging portfolios, the time horizon is therefore only one day and they are willing to accept such a setup if it can give accurate prices for that day.

Chapter 4

The Potential Approach for Interest and Exchange Rate Modelling

4.1 Introduction

As discussed in the review of interest rate modelling in Chapter 3, the majority of term structure models fall into one of two classes: those that model the spot rate process, or those that are concerned with the process of the forward rate. In this and subsequent chapters, we shall examine a relatively new approach to interest rate modelling, advocated by Constantinides [34] and Rogers [84], known as the potential approach.

The potential approach involves modelling interest rates by specifying the law of the state-price density process directly and makes use of the fact that the state-price density is a positive supermartingale. This modelling approach is still very much in its infancy and there appear to be very few papers related to it in the academic literature: Constantinides [34], Saa-Requejo [89], Rogers [84] and Rogers & Zane [85] are the only ones we know of. Moreover of these, Rogers & Zane [85], is the only one which investigates numerical fitting of the model to data and it is this area which will be the focus of our work in subsequent chapters.

The main concern of this chapter will be to introduce the background theory and fundamental results which we will rely on in our later work. Much of the theory presented in this chapter is taken from the work of Rogers [84].

4.2 Theoretical Foundations

We denote the time t price of a ZCB with maturity T $(t \leq T)$ by $P(t,T)$ and let $(r_t)_{t\geq 0}$ be the short-term interest rate process. Our starting point will be the expression given by the arbitrage pricing paradigm for the price of such a derivative (see §3.3.1),

$$P(t,T) \equiv \mathbb{E}_t[\exp(-\int_t^T r_s ds)], \tag{4.1}$$

where $\mathbb{E}_t \equiv \mathbb{E}[\cdot|\mathcal{F}_t]$ is the conditional expectation given the information $\mathcal{F}_t$ available at time t, taken with respect to a fixed risk-neutral measure. Suppose that there is some reference probability $\tilde{\mathbb{P}}$ (equivalent to the risk-neutral measure $\mathbb{P}$) and define the state-price density process $(\zeta_t)_{t\geq 0}$ by

$$\begin{aligned} \zeta_t &:= \exp(-\int_0^t r_s ds) \left.\frac{d\mathbb{P}}{d\tilde{\mathbb{P}}}\right|_{\mathcal{F}_t} \\ &:= \exp(-\int_o^t r_s ds) Z_t. \end{aligned} \tag{4.2}$$

Since we are concerned with nominal interest rates, $r \geq 0$, the process ζ_t is a positive supermartingale. If ζ also tends to 0 in L^1, that is,

$$\tilde{\mathbb{E}}\zeta_t \equiv P(0,t) \to 0 \text{ as } t \to \infty,$$

then ζ is called a potential. This terminology has strong links to the Markov process concept of a potential.[1]

The following lemma uses the definition of the state-price density given in (4.2) to rework the arbitrage price of a ZCB given in (4.1).

Lemma 4.1 *The time t price of a ZCB is*

$$P(t,T) = \tilde{\mathbb{E}}_t[\zeta_T]/\zeta_t \tag{4.3}$$

[1] see, for example, Rogers & Williams [11], Section II 54

Proof

$$
\begin{aligned}
P(t,T) &= \mathbb{E}_t[\exp(-\int_t^T r_s ds)] \\
&= \tilde{\mathbb{E}}_t\left[\exp(-\int_t^T r_s ds)\left.\frac{d\mathbb{P}}{d\tilde{\mathbb{P}}}\right|_{\mathcal{F}_T}\right] / \left.\frac{d\mathbb{P}}{d\tilde{\mathbb{P}}}\right|_{\mathcal{F}_t} \\
&= \tilde{\mathbb{E}}_t\left[\exp(-\int_0^T r_s ds)\exp(\int_0^t r_s ds)\left.\frac{d\mathbb{P}}{d\tilde{\mathbb{P}}}\right|_{\mathcal{F}_T}\right] / \left.\frac{d\mathbb{P}}{d\tilde{\mathbb{P}}}\right|_{\mathcal{F}_t} \\
&= \tilde{\mathbb{E}}_t[\zeta_T \exp(\int_0^t r_s ds)] / \left.\frac{d\mathbb{P}}{d\tilde{\mathbb{P}}}\right|_{\mathcal{F}_t} \\
&= \tilde{\mathbb{E}}_t[\zeta_T]\exp(\int_0^t r_s ds) / \left.\frac{d\mathbb{P}}{d\tilde{\mathbb{P}}}\right|_{\mathcal{F}_t} \\
&= \tilde{\mathbb{E}}_t[\zeta_T]/\zeta_t.
\end{aligned}
$$

□

This approach of representing the bond price in terms of the state-price density and then modelling this process, was first proposed by Constantinides [34]. Rogers [84] demonstrates how the theory of Markov processes can be used to generate examples of families of positive supermartingales thus giving a wide family of models for this approach.

Throughout our work on the potential model we consider a Markov process $(X_t)_{t\geq 0}$ with state space $\mathcal{X}$ and resolvent $(R_\lambda)_{\lambda\geq 0}$ and generate a family of supermartingales by setting

$$\zeta_t := e^{-\alpha t} R_\alpha g(X_t), \tag{4.4}$$

where $g : \mathcal{X} \to [0, \infty)$ and $\alpha > 0$.

This choice of supermartingale is given in generic approach 1 of Rogers [84] except that Rogers considers the normalised version of ζ_t so that $\zeta_0 = 1$.

To verify that (4.4) is indeed a positive-valued supermartingale, we consider the martingale

$$
\begin{aligned}
M_t &\equiv \mathbb{E}[\int_0^\infty e^{-\alpha s} g(X_s) ds | \mathcal{F}_t] \\
&= \mathbb{E}[\int_t^\infty e^{-\alpha s} g(X_s) ds + \int_0^t e^{-\alpha s} g(X_s) ds | \mathcal{F}_t] \\
&= \int_0^t e^{-\alpha s} g(X_s) ds + \mathbb{E}[\int_t^\infty e^{-\alpha s} g(X_s) ds | \mathcal{F}_t] \\
&= \int_0^t e^{-\alpha s} g(X_s) ds + \mathbb{E}[\int_0^\infty e^{-\alpha(t+u)} g(X_{t+u}) du | X_t] \\
&= \int_0^t e^{-\alpha s} g(X_s) ds + e^{-\alpha t} R_\alpha g(X_t). \qquad (4.5)
\end{aligned}
$$

Now since $A_t \equiv \int_0^t e^{-\alpha s} g(X_s) ds$ is an increasing process, $\zeta_t = M_t - A_t$ must be a supermartingale. Clearly ζ_t is positively valued for all t: by the positivity of the resolvent operator and since $g \geq 0$.

This specification for the potential provides us with a very useful family of positive supermartingales. Different choices of g and α give a wide range of state-price densities.

Next we shall derive an expression for the interest rate process. Differentiating (4.2) yields

$$
\begin{aligned}
d\zeta_t &= \exp(-\int_0^t r_s ds) dZ_t - Z_t \exp(-\int_0^t r_s ds) r_t dt \\
&= \zeta_t (Z_t^{-1} dZ_t - r_t dt). \qquad (4.6)
\end{aligned}
$$

Differentiating

$$
\zeta_t = M_t - \int_0^t e^{-\alpha s} g(X_s) ds,
$$

gives

$$
d\zeta_t = dM_t - e^{-\alpha t} g(X_t) dt,
$$

hence

$$\zeta_t^{-1} d\zeta_t \;=\; \zeta_t^{-1} dM_t - \frac{e^{-\alpha t} g(X_t)}{e^{-\alpha t} R_\alpha g(X_t)} dt.$$

Hence from the definition of Z_t as the change of measure,

$$Z_t \equiv \zeta_t \exp\left(\int_0^t r_s ds\right) \tag{4.7}$$

is a local martingale and we see that if we were using ζ as the state-price density relative to the probability space of the Markov property X, the spot rate process is

$$r_t \equiv \frac{g(X_t)}{R_\alpha g(X_t)}. \tag{4.8}$$

4.3 Modelling multiple countries

It follows from the last section that, to use the potential approach to model the term structure for a single country, we must first model a single Markov process and then construct a term structure model as a function of that Markov process.

However, what is remarkable about this approach is that in order to model the term structure for multiple countries, we still only require a single Markov process. We model the one Markov process and we can then add as many countries as we like by simply taking a new function of the Markov process for each new country. This fact is important and represents a major benefit the potential approach has over many other modelling regimes. Most models require new sources of randomness for each country modelled and this rapidly increases the computational and theoretical burden. Here, under the potential approach, all the randomness can be explained by one base Markov process.

Therefore, under the construction for the state-price density in the previous section, for each country i we require a function $g_i : \mathcal{X} \to [0, \infty)$ and scalar $\alpha_i > 0$.

4.4 Modelling Exchange Rates

One of the most exciting consequences of this approach is the ease with which exchange rates can be modelled in this framework. Saa-Requejo [89] and Rogers [84] both observed that under certain assumptions the exchange rate between two countries is the ratio of their state-price densities. Therefore, as we are essentially modelling the state-price density process for each country, it is clear that once the term structure for two countries has been modelled, the exchange rate between them will be determined for free; no new sources of randomness are required.

To see that the exchange rate between the two countries is indeed the ratio of the state-price densities of each country, we present the same argument as that given in the Appendix of Rogers [84]. This is a standard change of numeraire argument. The expression we shall give here differs slightly from that given in Rogers [84] because there it is assumed $\zeta_0^i = 1$ for all countries i.

We consider the situation where we have several countries and the assets in one country are freely exchangeable, at the prevailing exchange rate, for assets in another country.

Let $(S_t^i)_{t\geq 0}$ be the price process for a non-dividend paying asset in country i, measured in the currency of country i. Then if ζ_t^i is the state-price density for country i's assets, $(\zeta_t^i S_t^i)_{t\geq 0}$ is a $\tilde{\mathbb{P}}$-local martingale.

Moreover, assuming complete markets, ζ^i is (up to multiples) the only process for which $\zeta^i S$ is a $\tilde{\mathbb{P}}$-local martingale for every traded asset S.

Next, denoting C_t^{ij} to be the time-t value of 1 unit of country j's currency in country i's currency, observe that

$(C_t^{ij} S_t^j)_{t\geq 0}$ is the price process in country i's currency of the asset whose price process in country j is $(S_t^j)_{t\geq 0}$.

From this it follows that

$$\zeta_t^i C_t^{ij} S_t^j \text{ is a } \tilde{\mathbb{P}}\text{-local martingale,}$$

however, we also know that

$$\zeta_t^j S_t^j \text{ is a } \tilde{\mathbb{P}}\text{-local martingale.}$$

Therefore, in the complete situation we conclude that

$$C_t^{ij} \frac{\zeta_t^i}{\zeta_t^j} = C_0^{ij} \frac{\zeta_0^i}{\zeta_0^j}. \tag{4.9}$$

4.5 Existing numerical investigations

As was mentioned at the beginning of this chapter, Rogers & Zane [85] is the only paper we know of which considers the fitting of potential models to interest and exchange rate data. Rogers and Zane investigate the fit of one class of potential models to daily yield curves for the US dollar and sterling. They present results for modelling the exchange rates between the two currencies together with the yield curves for each country over a period of 200 days from January 1991.

For the underlying Markov process, Rogers and Zane consider a two dimensional Gaussian diffusion X solving

$$dX_t = dW_t - BX_t dt \tag{4.10}$$

where B is a 2×2 matrix and W_t is a Brownian Motion in $\mathbb{R}^2$. The bond price under this Markov process is an expectation of a function of X. The distribution of X_T is $N(e^{-TB}X_0, V_T)$ where

$$V_T = \int_0^T e^{-sB}(e^{-sB})^T ds. \tag{4.11}$$

This choice of diffusion process is appealing because of its tractability; prices of bonds and other derivatives can be reduced to an integration with respect to a Gaussian density and Rogers [84] gives closed form expressions for this case.

The numerical work in Rogers & Zane [85] is restricted to Brownian motions in two dimensions because of the increased computational burden involved in considering higher order diffusions.

4.6 Conclusions

The potential approach to term structure modelling has many appealing features, one of these is the ease with which the yield curves and exchange rates for several countries can be handled at once. Most models require additional sources of randomness for every country and exchange rate introduced into the model. This is not necessary with potential models; by modelling a single Markov process we can add as many countries as we wish. Each yield curve is modelled via a function of the Markov process and when a new country is added, all that is required is to add a new function of the Markov process.

Even with this alternative approach to modelling interest rates, a number of the traditional issues still remain. For example, the question of a suitable chöice for the underlying Markov process.

In Chapters 5 and 6 we shall examine the potential approach further, constructing a framework under which we can calibrate a potential model where the underlying process is a Markov chain.

Chapter 5

An Investigation into Markov Chain Potential Models

5.1 Introduction, Aims and Objectives

In this and the next chapter we will investigate potential models whose underlying structure is a finite Markov chain. As discussed in the conclusions of the last chapter, the choice of the underlying process is a key question for any interest rate model. Most models are forced to choose an underlying diffusion primarily on the basis of tractability and not always because of any realistic behaviour the diffusion may exhibit.

One of the major advantages of the potential approach is that it is not necessary to restrict oneself to the situation where the underlying Markov process is a diffusion. We shall explore letting the Markov process in our potential model be a data-imposed homogeneous time-continuous finite state Markov chain. Markov chain interest rate models have been suggested before but they have never been investigated within this approach.

Inevitably, the emphasis of much of this study will be numerical. The aim will be to present a logical advancement and development of a framework in which potential Markov chain models can be calibrated to accurately price interest rate derivatives. Therefore, we will present a stage-by-stage analysis documenting the results of the trials carried out and discussing the benefits and flaws of each approach. Where possible we will propose improvements of the initial trials and investigate their effect. Throughout the investigation emphasis will be placed on creating computationally feasible models that could be adopted in a trading environment.

The objective of this study **will not** be to attempt to conclusively determine whether the Markov chain potential approach is demonstrably better than other interest rate models. Our objective is more to investigate this new approach and report on the best method of implementation which we have found. Of course, we intend to give a number of results; however, because of the lack of a common general testing framework for interest rate models it is difficult to judge the quality of our fits with those reported for other models. Therefore, the work here will not serve as an empirical comparison of interest rate models and we will not make any categorical statements regarding the relative performance of the various models in the literature.

The basic construction of this chapter is as follows. We begin in §5.2 and §5.3 by introducing some definitions and discussing the specific details of the chains we shall consider. In these sections we also derive explicit formulae for the price of ZCB's, caps and swaps within this framework, hopefully demonstrating the flexibility of the approach and the ease with which these prices can be obtained.

In §5.4 we present a theoretical result showing how it is possible to use a potential model to replicate the bond prices of the Vasicek spot rate model. We then demonstrate how this can be implemented at a practical level. This involves constructing a generator matrix for a chain exhibiting the behaviour of a mean reverting Ornstein Uhlenbeck process and we make use of the formulae derived in §3.3.3 and §5.3.

5.2 Continuous-time Markov Chain Technology

There are a number of appealing reasons for looking at Markov chains instead of diffusion based models. Traditionally academics and practitioners have often had to compromise between a diffusion which exhibits realistic properties and one which is tractable enough to provide straightforward expressions for the various popular interest rate derivatives (IRD's) which they must price. The calculation of prices for instruments such as swaps, caps, floors and other IRD's may often be reduced to an integration over the statespace but these prices can only be given in closed form for certain diffusions. The absence of explicit expressions can often imply that numerical integration is the only solution and this can be cumbersome as well as computationally time consuming, thus making it more likely to be rejected by the practitioners.

These problems are not encountered when dealing with finite state Markov chains; we no longer have an integration over the statespace just a sum over the states. Moreover, adding states is only adding extra terms to the sum and therefore from a computational point of view this is incredibly appealing.

There are several Markov chain structures we could use. One structure we discuss in forthcoming sections is a discrete version of the mean reverting Ornstein Uhlenbeck process, that is, a random walk which fluctuates around some mean value and which is always pulled back towards the mean by a force proportional to the current deviation from the mean. This walk may be on a circle, a torus or some other network. Other examples of structures which could be adopted are ones which exhibit cyclical movements or even purely erratic behaviour. The options are limitless and the flexibility is very appealing.

5.3 Explicit Formulae for Prices of ZCB's, Caps and Swaps

In order to demonstrate how straightforward it is to derive pricing formulae in the Markov chain potential framework, we shall derive explicit expressions for the price of three main contracts, the ZCB, the caplet and the plain vanilla swaption. We begin by introducing our notation.

Let $(X(t))_{t\in\mathbb{R}^+}$ be a continuous-time Markov chain with finite countable statespace $\mathcal{X}$. We let $X(t) = j$, (for some $j \in \mathcal{X}$) denote the fact that the process is in state j at time t. Thus, $X(t)$ is referred to as the state of the process X at time t. We may use the alternative notation $X_t = X(t)$ when there is no ambiguity.

We denote by Q the infinitesimal generator (or Q-matrix) of the finite Markov chain X and $(P(t))_{t\geq 0}$ is the transition semigroup (of matrices). This semigroup is expressible in terms of Q as $P(t) = \exp(tQ)$.

Throughout this study we define the state-price density, ζ, as

$$\zeta_t = e^{-\alpha t} R_\alpha g(X_t). \tag{5.1}$$

5.3.1 Pricing a ZCB

Starting from the result of lemma 5.1, we see that

$$\begin{aligned} P(t,T) &= \tilde{\mathbb{E}}_t[\zeta_T]/\zeta_t \\ &= \sum_{k\in\mathcal{X}}[P_{T-t}]_{ik}\frac{e^{-\alpha T}R_\alpha g(k)}{e^{-\alpha t}R_\alpha g(i)} \\ &= \sum_{k\in\mathcal{X}}[e^{(T-t)Q}]_{ik}e^{-\alpha(T-t)}\left[\frac{R_\alpha g(k)}{R_\alpha g(i)}\right] \end{aligned} \tag{5.2}$$

where $X_t = i$ and $\tilde{\mathbb{E}}$ is the expectation taken with respect to a reference measure.

Therefore, given a generator matrix Q of some underlying Markov structure, together with a non-negative function g, the scalar α and the start state of the chain, we can compute the price of a bond.

5.3.2 Pricing a Caplet

An interest rate cap is a contractual agreement which effectively guarantees to its holder an upper limit for the variable rate on some loan. Hence it protects the holder against floating interest rates being too high. In practice, the seller of an interest rate cap agrees to pay cash to its holder if the variable interest rate on some loan exceeds a mutually agreed fixed level at some future date or dates. The holder is not affected by the agreement if the variable rate is ultimately more favourable to them than the agreed fixed level.

A typical contract may be as follows. Let us denote by K the fixed interest rate offered by an interest rate cap on some principal value L. The agreement commences on date T_0 and the cap is settled in arrears on dates $T_i = T_0 + i\delta$ $(i = 1,..,n)$. We will denote the period-δ LIBOR rate at time T_j by $R(T_j)$ and we let r_s denote the instantaneous short term interest rate. Payments continue for the lifetime of the option and each of the individual cash-flows is called a caplet. Thus a cap is the sum of many caplets, so if we find an expression for the price of a caplet we get the expression for the corresponding cap for free.

Consider a caplet with start time T_{n-1} and settlement $T_n = T_{n-1} + \delta$. The cash-flow at time T_n is

$$(R(T_{n-1}) - K)^+ L\delta.$$

The value of the caplet at time $t < T_{n-1}$ is

$$\begin{aligned}
\mathrm{Cpl}(t, T_{n-1}, T_n) &= \mathbb{E}_t\left[\exp(-\int_t^{T_{n-1}} r_s ds) P(T_{n-1}, T_n) L\delta (R(T_{n-1}) - K)^+\right] \\
&= \tilde{\mathbb{E}}_t\left[\frac{\zeta_{T_{n-1}}}{\zeta_t} P(T_{n-1}, T_n) L\delta (R(T_{n-1}) - K)^+\right] \qquad (5.3)
\end{aligned}$$

where,

- $\mathbb{E}_t \equiv \mathbb{E}[\cdot|\mathcal{F}_t]$ is the conditional expectation given the information $\mathcal{F}_t$ available at time t, taken with respect to a fixed risk-neutral measure;
- $\tilde{\mathbb{E}}_t \equiv \tilde{\mathbb{E}}[\cdot|\mathcal{F}_t]$ is again a conditional expectation but this time with respect to a reference measure.

Now since $P(T_{n-1}, T_n) = (1 + R(T_{n-1})\delta)^{-1}$ from (5.3) we get

$$\begin{aligned}
\mathrm{Cpl}(t, T_{n-1}, T_n) &= \tilde{\mathbb{E}}_t\left[\frac{\zeta_{T_{n-1}}}{\zeta_t} P(T_{n-1}, T_n) L\delta \left\{(P(T_{n-1}, T_n)^{-1} - 1)\delta^{-1} - K\right\}^+\right] \\
&= \tilde{\mathbb{E}}_t\left[\frac{\zeta_{T_{n-1}}}{\zeta_t} L \left\{1 - P(T_{n-1}, T_n) - K\delta P(T_{n-1}, T_n)\right\}^+\right] \\
&= \tilde{\mathbb{E}}_t\left[\frac{\zeta_{T_{n-1}}}{\zeta_t} L \left(\frac{1}{1 + K\delta} - P(T_{n-1}, T_n)\right)^+\right](1 + K\delta) \\
&\equiv \tilde{\mathbb{E}}_t\left[\frac{\zeta_{T_{n-1}}}{\zeta_t} L \left(\tilde{K} - P(T_{n-1}, T_n)\right)^+\right](1 + K\delta). \qquad (5.4)
\end{aligned}$$

Therefore, we see that a caplet is a put option on a ZCB, so we have reduced the problem of pricing a caplet to that of pricing an option on a ZCB. Taking $X_t = i$ we can proceed;

$$\begin{aligned}
\mathrm{Cpl}(t, T_{n-1}, T_n) &= \tilde{\mathbb{E}}_t\left[\frac{\zeta_{T_{n-1}}}{\zeta_t} L \left\{\tilde{K} - P(T_{n-1}, T_n)\right\}^+\right](1 + K\delta) \\
&= \sum_{k \in \mathcal{X}} [e^{(T_{n-1} - t)Q}]_{ik} \left(\frac{e^{-\alpha T_{n-1}} R_\alpha g(k)}{e^{-\alpha t} R_\alpha g(i)}\right) L \left\{\tilde{K} - P(T_{n-1}, T_n)\right\}^+ (1 + K\delta)
\end{aligned}$$

and this expression can be computed with the aid of (5.2). As observed earlier, it is a trivial extension to price a cap.

5.3.3 Pricing a Swaption

In this subsection we shall consider pricing a plain vanilla interest rate swaption. There are a large variety of swap agreements available and further information on these can be found in Hull [4] and Musiela & Rutkowski [7] as well as within the specialist literature.

A plain vanilla swap is an agreement between two parties to exchange future cash-flows on some future prearranged dates. One side, party A, agrees to pay the other side, party B, a fixed interest rate on some notional principal. In return, party B agrees to pay party A interest at a floating rate on the same principal for the same period of time.

There are essentially two perspectives from which we could attempt to price a swap. If we take the view of the party which pays the fixed rate and receives floating then we call it a **payer swap**. The converse situation is called a **receiver swap**.

We will consider a payer swap. The payments will be settled in arrears so that the floating rate is determined at the beginning of each time period and paid at the end of it. The fixed rate of interest is usually chosen so that the value of the swap at the time it is entered into is zero.

We denote the notional principal by L and the fixed interest rate by K. The agreement starts at date T_0 with payments in arrears at dates $T_i = T_0 + i\delta$ $(i = 1, .., N)$. Therefore, the fixed interest payments are $L\delta K$ at each of the payment dates. The time-t $(t \leq T_0)$ total value of all fixed interest payments is

$$L\delta K \sum_{j=1}^{N} P(t, T_j). \tag{5.5}$$

Next, we must obtain an expression for the time-t total value of the floating payments. First observe that the payment for one floating leg of a swap settling at date T_j, say, $(j \in 1, .., N)$ is exactly equal to a deposit of L at date T_{j-1} and a withdrawal of L at date T_j. So this is equivalent to two bonds. From this it follows that the total value of the floating side payments of our swap at time-t must be

$$(P(t, T_0) - P(t, T_N))L. \tag{5.6}$$

Therefore, (5.5) and (5.6) imply that for the swap to have zero initial value the fixed rate of interest, K, must be

$$K = \frac{1 - P(T_0, T_N)}{\delta \sum_{j=1}^{N} P(T_0, T_j)}. \tag{5.7}$$

We are now in a position to price a swaption. A swaption is an option to enter into a swap agreement on a future date at a given rate. Suppose we have a swaption to pay a fixed rate K on a swap which starts at date T_0 with payment dates $T_i = T_0 + i\delta$ ($i = 1, .., N$), then the payoff of the option at time T_0 is

$$\left(\frac{1 - P(T_0, T_N)}{\delta \sum_{j=1}^{N} P(T_0, T_j)} - K \right)^+ . \tag{5.8}$$

Hence the time-t value will be

$$\text{Swp}(t, T_0) = \mathbb{E}_t \left[\exp(-\int_t^{T_0} r_s ds) \left\{ \frac{1 - P(T_0, T_N)}{\delta \sum_{j=1}^{N} P(T_0, T_j)} - K \right\}^+ \right]$$

and from this we can obtain a potential model pricing expression, taking $X(t) = i$,

$$\begin{aligned} \text{Swp}(t, T_0) &= \mathbb{E}_t \left[\exp(-\int_t^{T_0} r_s ds) \left\{ \frac{1 - P(T_0, T_N)}{\delta \sum_{j=1}^{N} P(T_0, T_j)} - K \right\}^+ \right] \\ &= \tilde{\mathbb{E}}_t \left[\frac{\zeta_{T_0}}{\zeta_t} \left(\frac{1 - P(T_0, T_N)}{\delta \sum_{j=1}^{N} P(T_0, T_j)} - K \right)^+ \right] \\ &= \sum_{k \in \mathcal{X}} [P_{(T_0 - t)}]_{ik} \left(\frac{e^{-T_0 \alpha} R_\alpha g(k)}{e^{-t\alpha} R_\alpha g(i)} \right) \psi(i, T_0, T_N), \end{aligned}$$

where

$$\psi(X(t), T_0, T_N) = \left(\frac{1 - P(T_0, T_N)}{\delta \sum_{j=1}^{N} P(T_0, T_j)} - K \right)^+$$

So our final expression is

$$\text{Swp}(t, T_0) = \sum_{k \in \mathcal{X}} [e^{(T_0 - t)Q}]_{ik} e^{-(T_0 - t)\alpha} \left(\frac{R_\alpha g(k)}{R_\alpha g(i)} \right) \psi(i, T_0, T_N).$$

5.4 Replicating Vasicek Prices Via the Potential Approach

In some cases it is possible to create potential models which precisely replicate the dynamics of particular spot rate models. As an example, in this section we will generate a potential model which replicates the prices of the Vasicek spot rate model. We present first the theoretical setup and then demonstrate how this can be implemented numerically.

Consider Vasicek's model in the form

$$dr_t = \sigma dW_t + \beta(\rho - r_t)dt \tag{5.9}$$

where ρ, β, σ are fixed positive constants and W_t is a standard 1-d Brownian motion in the risk-neutral measure. We have already examined this equation in §3.3.3.

For us to create a potential model which has the same dynamics as (5.9), we need to choose an appropriate underlying diffusion to use for the potential model. The obvious choice for this is the Ornstein Uhlenbeck process because the Vasicek model is itself just a mean reverting version of this process.

Therefore, let's consider a potential model with underlying Markov process

$$dX_t = \tilde{\sigma} d\tilde{W}_t - \tilde{\beta} X_t dt \tag{5.10}$$

where $\tilde{\sigma}$, $\tilde{\beta}$ are fixed. **Importantly**, observe that $\tilde{W}_t$ is a 1-d Brownian motion in the **reference** measure, $\tilde{\mathbb{P}}$.

Next, we define g so that $R_{\tilde{\alpha}} g(X_t) = \exp(X_t/\tilde{\beta})$ and therefore using (5.1) the state-price density is

$$\zeta_t = e^{-\tilde{\alpha} t} \exp\left(\frac{X_t}{\tilde{\beta}}\right), \tag{5.11}$$

where $\tilde{\alpha} > 0$. We can obtain an expression for the interest rate by using Itô's formula,

$$d\zeta_t = \left(-\tilde{\alpha}\zeta_t - X_t\zeta_t + \frac{\tilde{\sigma}^2}{2\tilde{\beta}^2}\zeta_t\right) dt + \frac{\tilde{\sigma}}{\tilde{\beta}}\zeta_t d\tilde{W}_t.$$

Therefore,

$$\begin{aligned}\frac{d\zeta_t}{\zeta_t} &= -\left(\tilde{\alpha} + X_t - \frac{\tilde{\sigma}^2}{2\tilde{\beta}^2}\right) dt + \frac{\tilde{\sigma}}{\tilde{\beta}} d\tilde{W}_t \\ &\equiv -r_t dt + \frac{\tilde{\sigma}}{\tilde{\beta}} d\tilde{W}_t.\end{aligned}$$

Hence

$$r_t = \left(\tilde{\alpha} + X_t - \frac{\tilde{\sigma}^2}{2\tilde{\beta}^2}\right). \tag{5.12}$$

The Radon-Nikodym derivative is

$$Z_t = \left.\frac{d\mathbb{P}}{d\tilde{\mathbb{P}}}\right|_{\mathcal{F}_t} = \exp\left(\frac{\tilde{\sigma}}{\tilde{\beta}}\tilde{W}_t - \frac{\tilde{\sigma}^2}{2\tilde{\beta}^2}t\right)$$

and

$$W_t = \tilde{W}_t - \left(\frac{\tilde{\sigma}}{\tilde{\beta}}\right) t.$$

Therefore, rewriting (5.10) in terms of a Brownian motion with respect to the risk-neutral measure gives

$$dX_t = \tilde{\sigma} dW_t + \tilde{\beta}\left(\frac{\tilde{\sigma}^2}{\tilde{\beta}^2} - X_t\right) dt. \tag{5.13}$$

From (5.12) we have $dr_t = dX_t$, so that (5.13) becomes

$$dr_t = \tilde{\sigma} dW_t + \tilde{\beta}(\frac{\tilde{\sigma}^2}{2\tilde{\beta}^2} + \tilde{\alpha} - r_t)dt. \tag{5.14}$$

Finally comparing (5.14) with Vasicek's model (5.9) we see that to replicate the dynamics of the Vasicek spot rate model the parameters of the potential model are given in terms of the parameters of Vasicek's model as follows:

$$\begin{aligned}\tilde{\alpha} &= \rho - \frac{\sigma^2}{2\beta^2} \\ \tilde{\sigma} &= \sigma \\ \tilde{\beta} &= \beta.\end{aligned} \tag{5.15}$$

Therefore to summarise, we have found that by taking the underlying diffusion in the potential approach to be an Ornstein Uhlenbeck process and setting the potential parameters as in (5.15), we can exactly replicate the prices given by Vasicek. Our aim now is to examine how this can be implemented numerically within a Markov chain potential model and to compare bond prices for this model with those of Vasicek in order to verify that they are in fact close to each other. The level of closeness will depend on how good our approximation of the Ornstein Uhlenbeck process is, to improve accuracy we can take more states in the Markov chain.

5.4.1 Markov Chain Implementation

The first task is to construct a Markov chain approximation for the 1-d Ornstein Uhlenbeck process

$$dX_t = \tilde{\sigma} d\tilde{W}_t - \tilde{\beta} X_t dt. \tag{5.16}$$

This process takes the entire real line as its statespace, $\mathcal{X} = (-\infty, \infty)$ and has limiting distribution $\mathrm{N}(0, \tilde{\sigma}^2/(2\tilde{\beta}))$. For the approximation we shall discretise the $\mathbb{R}$ line by choosing states so that

$$x_j = \frac{\tilde{\sigma}}{\sqrt{2\tilde{\beta}}} \Phi^{-1}\left(\frac{j + \frac{1}{2} + M}{2M+1}\right)$$

for $j = -M, .., 0, .., M$ for some fixed $M \in \mathbb{N}$. Thus the total number of states is $(2M+1)$ and the value of M can be increased to improve accuracy.

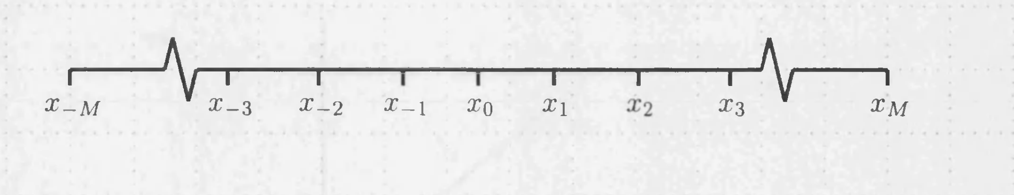

Figure 5-1: Discretisation of $\mathbb{R}$ line

We restrict the movements of the chain to immediately adjacent states, so that if the current state is x_j for $j \in \{-M+1, .., M-1\}$ then the next move must be to either x_{j-1} or x_{j+1}. The rate at which we move from state x_j to x_{j+1} will be denoted λ_j and we let μ_j denote the rate at which we move from state x_j to x_{j-1}. When the walk reaches an end state, the next jump must be back to the state from which the walk arrived. τ is the first jump time.

To obtain explicit expressions for the rates of change λ_j, μ_j, we shall use the fact that $e^{\tilde{\beta}t}X_t$ and $e^{2\tilde{\beta}t}(X_t^2 - \frac{\tilde{\sigma}^2}{2\tilde{\beta}})$ are martingales with (5.16). It is straightforward to verify that these are martingales by considering the Itô derivatives,

$$\begin{aligned} d(e^{\tilde{\beta}t}X_t) &= \tilde{\beta}(e^{\tilde{\beta}t}X_t)dt + e^{\tilde{\beta}t}dX_t + 0 \\ &= \tilde{\beta}(e^{\tilde{\beta}t}X_t)dt + e^{\tilde{\beta}t}\tilde{\sigma}dW_t - e^{\tilde{\beta}t}\tilde{\beta}X_t dt \\ &= e^{\tilde{\beta}t}\tilde{\sigma}dW_t \end{aligned}$$

and

$$\begin{aligned} d(e^{2\tilde{\beta}t}(X_t^2 - \frac{\tilde{\sigma}^2}{2\tilde{\beta}})) &= 2\beta e^{2\tilde{\beta}t}(X_t^2 - \frac{\tilde{\sigma}^2}{2\tilde{\beta}})dt + 2e^{2\tilde{\beta}t}X_t dX_t + e^{2\tilde{\beta}t}d < X >_t \\ &= 2\beta e^{2\tilde{\beta}t}(X_t^2 - \frac{\tilde{\sigma}^2}{2\tilde{\beta}})dt - 2\tilde{\beta}e^{2\tilde{\beta}t}X_t^2 dt + 2\tilde{\sigma}e^{2\tilde{\beta}t}X_t dW_t + \tilde{\sigma}^2 e^{2\tilde{\beta}t}dt \\ &= 2\tilde{\sigma}e^{2\tilde{\beta}t}X_t dW_t. \end{aligned}$$

Using the first martingale we have that for $j \in \{-M+1, .., M-1\}$

$$\begin{aligned} \mathbb{E}\left[e^{\tilde{\beta}\tau}X_\tau | X_0 = x_j\right] &= \int_0^\infty e^{\tilde{\beta}t}\left(x_{j+1}\left(\frac{\lambda_j}{\lambda_j+\mu_j}\right) + x_{j-1}\left(\frac{\mu_j}{\lambda_j+\mu_j}\right)\right)(\lambda_j+\mu_j)e^{-(\lambda_j+\mu_j)t}dt \\ &= \left[x_{j+1}\left(\frac{\lambda_j}{\lambda_j+\mu_j}\right) + x_{j-1}\left(\frac{\mu_j}{\lambda_j+\mu_j}\right)\right]\left(\frac{\lambda_j+\mu_j}{\lambda_j+\mu_j-\tilde{\beta}}\right) \\ &\equiv x_j. \end{aligned}$$

Therefore

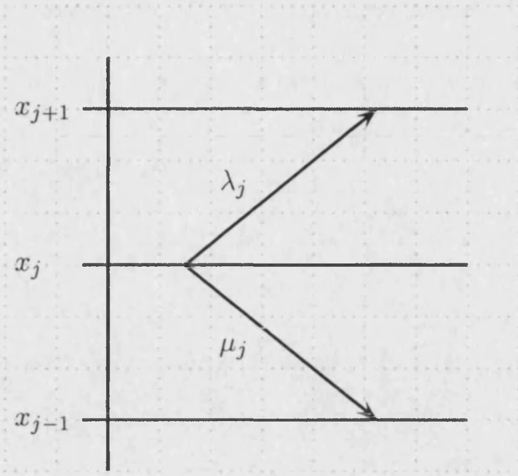

Figure 5-2: Jump Rates

$$(\lambda_j + \mu_j - \tilde{\beta})x_j = \lambda_j x_{j+1} + \mu_j x_{j-1}. \tag{5.17}$$

Next, using the second martingale

$$\begin{aligned}
\mathbb{E}\left[e^{2\tilde{\beta}\tau}(X_\tau^2 - \frac{\tilde{\sigma}^2}{2\tilde{\beta}})|X_0 = x_j\right] &= \int_0^\infty e^{2\tilde{\beta}t}\left((x_{j+1}^2 - \frac{\tilde{\sigma}^2}{2\tilde{\beta}})\lambda_j + (x_{j-1}^2 - \frac{\tilde{\sigma}^2}{2\tilde{\beta}})\mu_j\right)e^{-(\lambda_j+\mu_j)t}dt \\
&= \left[(x_{j+1}^2 - \frac{\tilde{\sigma}^2}{2\tilde{\beta}})\lambda_j + (x_{j-1}^2 - \frac{\tilde{\sigma}^2}{2\tilde{\beta}})\mu_j\right]\frac{1}{(\lambda_j + \mu_j - 2\tilde{\beta})} \\
&\equiv \left(x_j^2 - \frac{\tilde{\sigma}^2}{2\tilde{\beta}}\right).
\end{aligned}$$

Which yields

$$\lambda_j x_{j+1}^2 + \mu_j x_{j-1}^2 = x_j^2(\lambda_j + \mu_j - 2\tilde{\beta}) + \tilde{\sigma}^2. \tag{5.18}$$

Therefore, we have two equations (5.17) and (5.18) which we can solve for λ_j and μ_j when $j \in \{-M+1, .., M-1\}$ and so we have all the jump rates for the chain except for those at the end points. The end points must be dealt with separately; we only require one equation here so we solve

$$(\mu_M - \tilde{\beta})x_M = \mu_M x_{M-1}$$

for μ_M. Symmetry then gives us the value of λ_{-M}. By construction λ_M and μ_{-M} are both zero.

We are now in a position to form the generator matrix, Q, for our Ornstein Uhlenbeck Markov chain approximation.

5.4.2 Numerical Implementation

What remains is to put everything together and compare the bond prices given by the Vasicek spot rate model with those generated by our specially chosen potential Markov chain model.

As discussed in §3.3.3, the explicit pricing formula for ZCB's in the Vasicek model is given by

$$\begin{aligned} P(0,t) = \mathbb{E}[\exp(-\int_0^t r_u du)] &= \frac{1}{v_t\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(x-m_t)}{2v_t^2}} e^{-x} dx \\ &= \exp(-m_t + \frac{1}{2}v_t^2), \end{aligned} \tag{5.19}$$

where,

$$\begin{aligned} m_t &= \rho t + \frac{(1-\exp(-\beta t))}{\beta}(r_0 - \rho) \\ v_t^2 &= \frac{\sigma^2}{2\beta^3}[2\beta t - 3 + 4e^{-\beta t} - e^{-2\beta t}]. \end{aligned}$$

For the potential model, we use the specification given in §5.4.1 to form our generator matrix Q. Also, from (5.2) the price of a bond in this framework is given by

$$\sum_{k\in\mathcal{X}} [e^{(T-t)Q}]_{ik} e^{-\alpha(T-t)} \left[\frac{R_\alpha g(k)}{R_\alpha g(i)}\right]$$

and we set $R_\alpha g(X_t) = \exp(\frac{X_t}{\beta})$.

When implementing this numerically, we choose values for the Vasicek parameters, ρ, β and σ and then use the relations (5.15) to obtain the values for the potential parameters $\tilde{\alpha}$, $\tilde{\beta}$ and $\tilde{\sigma}$. The results showed that even with just a 15 state Markov chain, the two approaches gave the same price to within 6 significant figures.

5.5 Remarks

Proponents of forward rate models often argue that their framework is capable of replicating models that exist in the spot rate modelling paradigm. In this section it has been seen that the potential approach is capable of replicating the prices of a spot rate model.

The work here demonstrates that the potential approach is flexible enough to generate the same prices as another well-known interest rate model. This example also helps clarify how Markov chain potential models could be implemented in practice.

Chapter 6

Calibrating Markov Chain Potential Models to Interest Rates and Exchange Rates

6.1 Calibrating Markov chain potential models to market data

This chapter will concentrate on establishing a method for calibrating Markov chain potential models to market data. We focus on computationally feasible models which could be employed in a trading environment.

6.1.1 Discussion of Market Data

The data which is used in this study is daily yield curve data covering the period from 2nd January 1992 to 1st March 1996.

For each day we have values of the yield of bonds with maturity 1 month, 3 months, 6 months, 1 year, 2 years, 5 years, 7 years and 10 years. We shall use daily yield curve data for three currencies; these are sterling (GBP), United States dollar (USD) and the Deutschemark (DEM)[1].

We also have daily exchange rate data between these three currencies, obtained from the United States Federal Reserve Data Exchange[2].

[1]We are grateful to Dr Simon Babbs for supplying the GBP and DEM data. The USD data was taken from the website http://www.stls.frb.org/fred/index.html

[2]See http://www.federalreserve.gov/releases/H10/hist

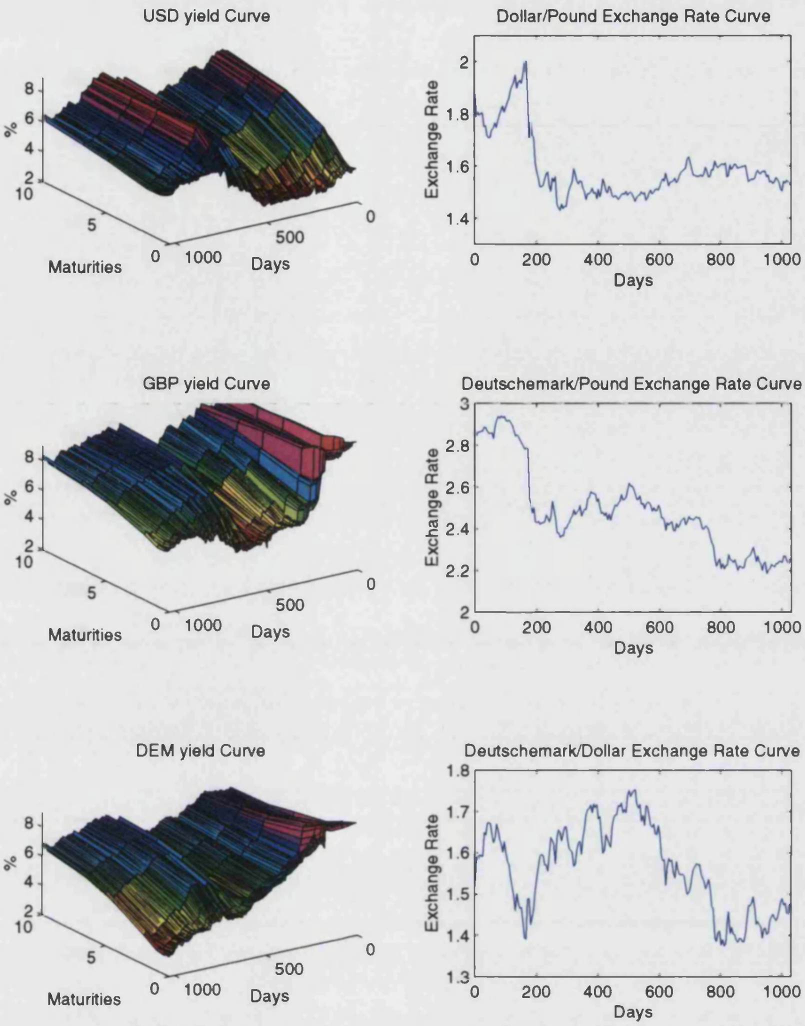

Figure 6-1: Yield and Exchange Rate Curves

The data sets were cleaned and any dates that were not common to every set were removed from all sets. This included public holidays and other days where one or more of the three markets was closed. In total we have 1029 days of traded data. Surface plots of the yield curve for each country, together with graphs of the exchange rates are shown in Figure 6-1.

It is worth pointing out that the period of time under consideration in this study represented a turbulent time in the world markets. The years of 1992 and 1993 saw both the US and UK economies in the middle of deep recessions. Indeed, 1992 was a year of huge turmoil for the UK economy, it saw the shock re-election of the Conservative party for a third consecutive term of office. This was followed by the day termed "Black-

Wednesday" (16th September 1992) in which the UK was embarrassingly forced out of the ERM and in which the Bank of England lost 4 billion pounds trying to stop the GBP devaluing. Moreover, on this day, the UK government announced a 5% rise in the base rate taking the rate to 15% in a desperate attempt to stop the pound's value sliding. The turmoil in the UK economy at this point was partly attributed (by many analysts) to the strength and dominance of the German currency. In fact it can be seen that the German economy had a strong influence on most of the other major European economies at this time.

Conversely, 1994 to 1996 saw a weakening of the German dominance and a recovery in the UK and US economies. These countries slowly came out of their long recession and this is reflected in the shape of the yield curve and exchange rate over this period. We have therefore chosen quite a varied and turbulent period for the calibration exercise.

Money market interest rates are frequently influenced by exogenous parameters set by monetary authorities. Babbs & Webber [15] [16] provide several illustrative examples demonstrating the effects that some of these parameters have on the short term interest rate. One particular example compares movements in 3 month sterling LIBOR with the Bank of England band one stop rate (a rate which closely follows the UK base rate). Figures 6-2 and 6-3 clearly show a link between 1 month and 3 month LIBOR and the Bank of England band one rate.

It follows therefore that if we were able to model the base rate (which is only a jump process), we would already have a reasonable model for 1 month LIBOR. Indeed, Babbs & Webber argue that term structure interest rate models should take into account the jump behaviour caused by monetary authorities exercising control over domestic markets. Moreover, they suggest that diffusion models of the term structure such as Duffie & Kan [42], HJM or extended Vasicek cannot capture the observed jump-diffusion behaviour. They present a theoretical framework where the short term interest rate follows a jump diffusion process. Their arguments provide us with a further incentive for pursuing Markov chain potential models as these models would seem ideally suited to modelling the occasional jumps caused by such exogenously determined parameters. The traditional approach of focusing on the volatilities of various yields and rates may actually be concentrating on the *noise* in the system, and overlooking the *signal*.

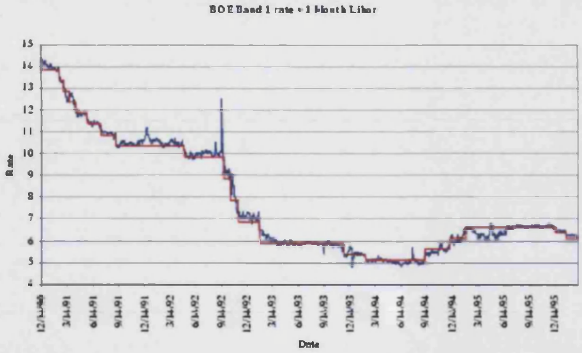

Figure 6-2: Base Rate against 1 Month LIBOR

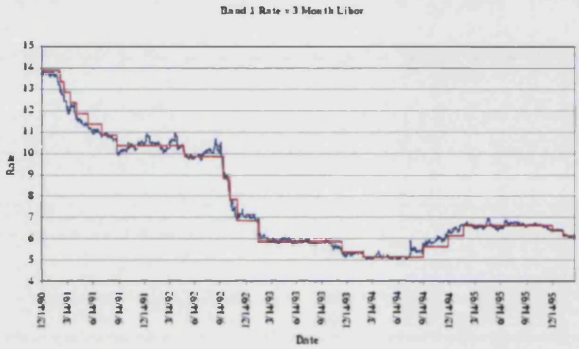

Figure 6-3: Base Rate against 3 Month LIBOR

6.1.2 Fitting Methodology

In §5.3.1 we saw that by taking

$$\zeta_t = e^{-\alpha t} R_\alpha g(X_t),$$

the time t price of a ZCB with maturity T is

$$P(t,T) = \sum_{k \in \mathcal{X}} [e^{(T-t)Q}]_{ik} e^{-\alpha(T-t)} \left[\frac{R_\alpha g(k)}{R_\alpha g(i)} \right].$$

In the context of a finite Markov chain X with generator matrix Q, the resolvent (or α-potential) is given by

$$R_\alpha = (\alpha I - Q)^{-1}.$$

The fundamental question now is how do we make these potential models generate accurate prices for interest rate derivatives? This question translates directly to what is the best way to find values for the parameters α and $\mathbf{g}$ and Q?

Before we explain the estimation methods we will use, let us introduce some notation. We shall denote the vector of parameters in our model by θ. On each day we will have market values for M observables and the market value on day n, for the ith observable will be denoted by y_n^i. Similarly, the model value for this observable will be denoted $Y^i(\theta, X_n)$ where X is the underlying Markov chain. We suppose that

$$y_n^i = Y^i(\theta, X_n) + \varepsilon_i \tag{6.1}$$

for $i = 1, .., M$, where ε_i represents independent Gaussian noise with $\mathbb{E}[\varepsilon_i] = 0$.

We adopt a Bayesian standpoint, and suppose that the initial law of X is given by $\pi = (\pi(i))_{i=1}^N$, where N is the number of states in the Markov chain. The initial law of θ is given by the density $f_0(\theta)$; conceptually θ is unchanging with time even though our knowledge of it varies.

We use the notation $\mathbf{z}_n \equiv (z_0, .., z_n)$. Based on the assumptions above, the likelihood Λ_n of $(\mathbf{X}_n, \mathbf{y}_n, \theta)$ is

$$\begin{aligned} \Lambda_n &\equiv \Lambda_n(\mathbf{X}_n, \mathbf{y}_n, \theta) \\ &= \frac{\pi(X_0) f_0(\theta)}{(2\pi|V|)^{n/2}} \prod_{j=1}^n p_{X_{j-1}X_j}(s_j; \theta) \exp[-b(y_j - Y(X_j; \theta))] \end{aligned} \tag{6.2}$$

where $p_{ij}(s, \theta) = \mathbb{P}_\theta(X_s = j | X_0 = i)$, and $b(z) \equiv \frac{1}{2} z.V^{-1} z$, where V is the covariance matrix of the Gaussian errors. $s_j = t_j - t_{j-1}$ is the time between the $(j-1)$th and jth observations. As we are interested in the posterior distribution of (X_n, θ) given $\mathbf{y}_n$, we introduce the notation

$$L_n(x, \mathbf{y}_n, \theta) = \sum_{\mathbf{X}_n : X_n = x} \Lambda_n(\mathbf{X}_n, \mathbf{y}_n, \theta), \tag{6.3}$$

and notice that,

$$L_n(x, \mathbf{y}_n, \theta) = \sum_{\xi} L_{n-1}(\xi, \mathbf{y}_{n-1}, \theta) p_{\xi x}(s_n; \theta) \exp[-b(y_n - Y(x; \theta))]. \tag{6.4}$$

For the Markov chain model in mind this expression will be far too complicated to allow exact analysis, and so we need to make simplifying assumptions to progress. These assumptions will be introduced as we progress.

The results of the fits will often be quoted in terms of the basis points (bp) error. A **basis point** is a percent of a percent and is a term used frequently in the industry and academic literature. We refer to

$$\text{basis point error per maturity} = \frac{1}{M} \sum_{i=1}^{M} 100 |y_n^i - Y^i(x; \theta)|. \tag{6.5}$$

6.1.3 Computational Considerations

In order to estimate the parameter vector θ, we will need to use a numerical minimisation routine. A number of routines were investigated and the best one (for our purposes) in terms of computational speed and numerical results was found to be the routine E04JYF supplied by NAG (Numerical Algorithms Group). Therefore this routine was used in all the trials reported in this study.

The routine itself is a quasi-Newton algorithm which can solve problems of the form

$$\text{Minimize } F(x_1, x_2, .., x_k) \tag{6.6}$$

subject to constant bounds l_j, u_j where $l_j \leq x_j \leq u_j$, $j = 1, 2, .., k$ when derivatives of $F(x)$ are unavailable. The number of iterations is dependent on the number of variables, the behaviour of $F(\mathbf{x})$ and the distance of the starting point from the solution.

At this stage we must also consider how we can compute the transition semigroup $\exp(tQ)$. It is important that this matrix can be computed rapidly and for this reason we restrict ourselves to Q matrices which are symmetrisable. That is, Q matrices which satisfy the time reversibility condition:

$$m_{ii}q_{ij} = m_{jj}q_{ji} \text{ for } i,j = 1,..,N \tag{6.7}$$

where $M = (m_{ij})$ is a diagonal matrix and $\sum_{i=1}^{N} m_{ii} = 1$.

Q is an $N \times N$ matrix and the diagonal matrix M plays the role of the **invariant mass** of the Markov chain. Observe that since $MQ = Q^T M$ we have

$$M^{\frac{1}{2}}QM^{-\frac{1}{2}} = M^{-\frac{1}{2}}Q^T M^{\frac{1}{2}}.$$

Hence, $M^{\frac{1}{2}}QM^{-\frac{1}{2}}$ is a real symmetric matrix from which it follows that Q is diagonalisable. Therefore, there exists an orthogonal matrix R such that

$$\Lambda = R^{-1}QR \tag{6.8}$$

is diagonal. Thus the semigroup is given by

$$e^{tQ} = Re^{t\Lambda}R^{-1}. \tag{6.9}$$

6.2 Size of the Markov chain statespace and period of calibration

The question we must address next is what is the optimum number of states for our Markov chain? To answer this, we use a technique from multivariate statistical analysis.

Our aim is to essentially develop taxonomies, that is, to examine the data and organise it into meaningful structures. This will help us to determine whether there are any clear groups of yield curve structures and if so how many there are. This information should give us a guideline for what would be a reasonable number of states for our Markov chain.

The method we shall use to investigate the taxonomy of the data will be a form of cluster analysis. Clearly, the absence of a-priori hypotheses about our data means that we are not able to directly use the traditional statistical significance tests. However, a cluster analysis is a viable alternative which will still allow us to detect significant natural homogeneous groupings in the data even without hypotheses. This technique is widely used in areas such as biology, medicine, and the social sciences to reveal relationships in large groups of data, see Hartigan [3].

The two key choices which need to be made when performing a cluster analysis are how to measure the 'distance' between objects, and then what method of grouping between objects to employ. The distances serve as a measure of similarity between objects and in our analysis we shall use the Euclidean metric distance. Euclidean distances are root sum-of-squares of differences. The criteria we use to link the variables will be the so called single linkage method. This is a nearest neighbour linkage which measures the distance to the nearest object within a group. We will perform the cluster analysis on a combined data set which includes all the GBP, USD and DEM data.

The results from the analysis are shown in Figures 6-4, 6-5 and 6-6. They indicate that there are between 11 and 14 major clusters. The dendrogram in Figure 6-6 perhaps shows this most clearly. However, a valid question would be where is an appropriate place to cut the tree because the lower you go in the dendrogram, the more clusters there are. One rule of thumb for this is to choose a place where the structure remains stable for long distances. Another possibility would be to look for cluster groupings that agree with existing or expected structures, or to replicate the analysis on subsets of the data to see if the structures emerge consistently. As determining long distances in a dendrogram is fairly subjective, we also analysed four randomly generated subsets from our data. Unfortunately, the results from these tests were inconclusive. This is slightly disturbing because it suggests that the clusters we have found here are specific to this particular data set.

An alternative way of determining how many states we should have in our Markov chain is to investigate fits using Markov chain potential models with states ranging from 8 to 50. Unfortunately, our computational limitations prevent us from carrying out this investigation. However, this computation task should not be a problem for the industry; if this approach were to be implemented, parallelization together with high specification computers would mean the computation time for Markov chains which have a large number of states would not be a significant concern.

Next we must decide which set of days we should use to test our calibration procedures. We consider fits over a period of 100 days, which is approximately a 4 month calendar period. As we have a sizeable dataset, we have choosen to split the 1029 days of data into 19 overlapping groups of 100 days, the statistics of which are shown in Table 6.1.

For the first calibration procedure we will present a summary of the results obtained when calibrating each of the 19 periods, then for more detailed analysis we choose to focus on period 14 and use this period for all future calibration exercises.

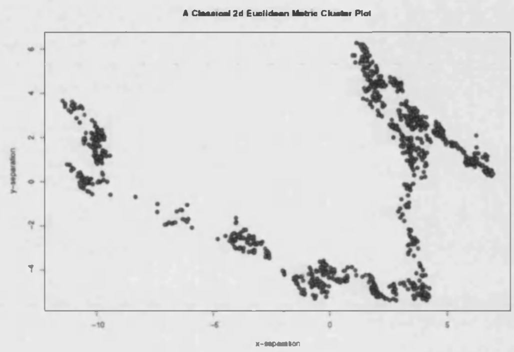

Figure 6-4: A classical 2d cluster plot generated using the mathematical statistics package Splus. This plot uses a Euclidean metric and single linkage method.

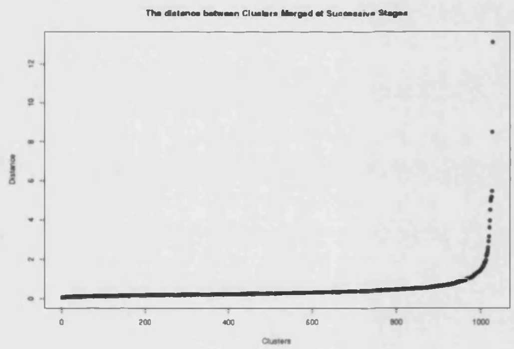

Figure 6-5: An alternative representation of the cluster analysis is where the distances between clusters merged at successive stages is represented. Produced using Splus.

Figure 6-6: A dendrogram which presents a plot of clusters at various distances. Produced using Splus.

	Calendar Period	Day Numbers	BRC	Daily change statistics in bp			
				Mean	Std Dev.	Min	Max
1	17th Feb 1992 - 15th July 1992	30-129	1	63.945	56.136	13.442	438.02
2	5th May 1992 - 25th Sept 1992	80-179	2	90.0	147.237	5.681	1139.932
3	16th July 1992 - 7th Dec 1992	130-229	3	125.785	166.37	5.681189	1139.932
4	28th Sept 1992 - 18th Feb 1993	180-279	3	98.061	107.264	10.073	668.876
5	8th Dec 1992 - 7th May 1993	230-329	1	65.253	55.419	7.840	525.249
6	19th Feb 1993- 19th July 1993	280-379	0	60.453	32.599	0.82144	157.463
7	10th May 1993 - 28th Sept 1993	330-429	0	59.916	37.846	0.82144	157.463
8	20th July 1993 - 8th Dec 1993	380-479	1	60.449	35.036	12.604	171.941
9	29th Sept 1993 - 23rd Feb 1994	430-529	2	75.179	68.243	12.604	560.769
10	9th Dec 1993 - 11th May 1994	480-579	1	84.796	69.671	12.698	560.769
11	24th Feb 1994 - 22 July 1994	530-629	0	74.533	37.415	9.360	196.289
12	12th May 1994 - 4th Oct 1994	580-679	1	74.245	46.334	9.360	278.727
13	25th July 1994 - 16th Dec 1994	630-729	2	77.238	47.230	13.700	278.727
14	5th Oct 1994 - 6th Mar 1995	680-779	2	83.205	45.743	13.700	217.905
15	19th Dec 1994 - 23rd May 1995	730-829	1	85.975	48.146	21.735	217.905
16	7th Mar 1995 - 3rd Aug 1995	780-879	0	82.916	50.674	6.466	207.79
17	24th May 1995 - 16th Oct 1995	830-929	0	60.324	46.884	6.466	207.79
18	4th Aug 1995 - 28th Dec 1995	880-979	1	37.773	21.0	5.597	124.308
19	17th Oct 1995 - 8th Mar 1996	930-1029	3	37.630	25.069	5.597	157.292

Table 6.1: Analysis of the GBP dataset for each of the 19 sample periods of 100 days. The column marked *mean* above refers to the average bp daily change over all maturities. The *Std Dev* is the standard deviation for the mean in bp. *Min*, *Max* refer to the bp change per day. *BRC* is used to denote the number of Bank of England base rate changes in the period.

6.3 Modelling Exercise 1: day-by-day calibration

The method of fitting we shall consider in this modelling exercise is a **day-by-day calibration**. We use this term to refer to a model which fits curves over a series of days with no constraints on parameter movements between days and no restriction on the parameters involved in the fit. Therefore, we choose to ignore all the 'earlier' information in (6.4) and use what is essentially a least squares method of estimation.

Given the observations y_n, on any day n, we simply compute

$$\min_{\theta} \left[(y_n - Y(x;\theta))^T V^{-1} (y_n - Y(x;\theta)) \right], \tag{6.10}$$

where in this minimisation we make the arbitrary convention that x is some distinguished state (the first say) in the statespace. The labelling of the states of the chain is irrelevant under this simplifying assumption. The matrix V^{-1} is chosen to be diagonal and can be modified so it represents the relative importance and significance of particular observables. For example, if we were keen to ensure that the model value for one particular observable is very accurate (i.e. very close to that given by the market) we would weight this observable higher than the others. For simplicity in this exercise, we only report results for the case where V is the identity.

The model values $Y(x;\theta)$ for each maturity are found by using the pricing formula for ZCB's given in (5.2) and then converting this to the yield to maturity value using definition 4.3,

$$Y(t,T) = \frac{-1}{T-t} \log P(t,T). \tag{6.11}$$

Although this calibration exercise cannot be expected to be very stable, it should provide us with a lower bound for the fitting error. Moreover, if the results of fitting under this assumption are disappointing, then the results will be disappointing under more realistic assumptions.

6.3.1 Choice of underlying structure

In §5.4 we presented a method for generating an approximation to the 1-d Ornstein Uhlenbeck mean reverting random walk on the $\mathbb{R}$ line. This is just one choice for the underlying structure and there are many others we could adopt: a walk on a circle, a random irreducible network, a mean reverting walk on a torus, or a cyclical structure are just some possibilities.

At the outset it is not clear what the best choice for the underlying structure is. The only immediate prerequisite is the need for the chain to be irreducible. That is, we require a chain with the property that no matter which state you are presently in, there is a positive probability of reaching all other states in finite time.

An alternative to choosing our own underlying structure is to allow the data to fix and impose the structure itself at the beginning of the fitting procedure. This approach has some immediate advantages

- if the underlying structure is influenced by the data we should have a model which understands and is able to adapt to the market more easily;
- we no longer have to concern ourselves with whether we have made the right choice for the Markov structure.

In this section we shall consider three different choices for the irreducible homogeneous Markov chain and investigate the results of each of these under a day-by-day calibration to market data.

6.3.2 Case A: a fixed Ornstein Uhlenbeck structure

To begin with we investigate day-by-day fits of a potential model where the chain is an approximation of a 1-d mean reverting Ornstein Uhlenbeck process on the real line.

We saw in §5.4.1 how to form a generator matrix representing this behaviour and we shall again make use of the procedure given in that section. However, in this case we will decide the parameter values $\tilde{\sigma}$ and $\tilde{\beta}$ at the start of the fitting procedure and these will remain fixed throughout the fit. In this sense, we are dealing with a fixed underlying structure and so the Q matrix and semigroups for each maturity need only be calculated once at the beginning of the fitting procedure. Computationally this is very desirable and efficient.

- *Method:* a day-by-day fit of the yield curve for one country with a fixed 11-state Markov structure.
- *Markov structure:* mean reverting Ornstein Uhlenbeck.
- *Data used:* GBP data only.

- *Parameters in the minimisation vector θ:*

$$\text{g vector and } \alpha.$$

- *Other settings:*

$$\tilde{\beta} = 0.2 \text{ and } \tilde{\sigma} = 0.2$$

6.3.3 Case B: a data imposed Ornstein Uhlenbeck structure

In this method we take the basic Ornstein Uhlenbeck mean reverting structure and instead of choosing the parameters $\tilde{\sigma}$ and $\tilde{\beta}$ at the beginning of the fitting process, we include them in the vector θ that we minimise over in (6.10). In this way, the data has an influence over the underlying structure.

- *Method:* a day-by-day fit of the yield curve for one country with a semi-data imposed 11-state Markov chain and the start state choosen to be the first state.
- *Data used:* GBP only.
- *Markov structure:* data imposed Ornstein Uhlenbeck.
- *Parameters in the minimisation vector θ:*

$$g \text{ vector, } \alpha, \text{ parameters } \tilde{\sigma},\ \tilde{\beta} \text{ relating to the underlying OU structure.}$$

6.3.4 Case C: an entirely data imposed underlying structure

Here we drop the mean reverting Ornstein Uhlenbeck process altogether and allow the market data to entirely choose the form of the underlying Markov chain.

The most obvious approach here would appear to be to allow the minimisation routine to choose all the off diagonal elements in the Q matrix as non-negative values and then to determine the diagonal elements so that the matrix is stochastic. Unfortunately, this approach causes some difficulties. To start with, we want our Markov chain to be irreducible and the above idea would not necessarily give irreducible chains!

Another concern would be that in order to compute the semigroup we need to be able to find the eigenvalues and eigenvectors of the generator matrix Q. A general real unsymmetric matrix is not necessarily diagonalisable and even if it was, it could have complex eigenvalues and eigenvectors. The complex eigenvalues/vectors do not make the problem inaccessible but they do increase computation time.

A further obstacle to allowing complete flexibility is that if we were to allow the minimisation to choose all the off diagonal parameters we would have $(N^2 - N)$ extra parameters in the minimisaton $(N = |\mathcal{X}|)$ and even for an 11-state chain this is a large number of parameters.

Therefore, to overcome these objections, as explained in §6.1.3, we restrict our attention to Q matrices which are symmetrisable.

Computational procedure: to implement this, we first generate a symmetric flux matrix A and an invariant mass vector π. We need to find values for $a_{ij} > 0$ for $i > j$ which is $N(N-1)/2$ parameters and also $N-1$ entries of π.

We choose the matrix A by setting $a_{ij} = a_{ji}$ and computing all the diagonal elements so that $\sum_{j=1}^{n} a_{ij} = 0$. The π vector is normalised so that $\sum_{i=1}^{n} \pi_i = 1$. The generator matrix Q can then be formed using

$$a_{ij} = \pi_i q_{ij} = \pi_j q_{ji} = a_{ji}.$$

Since a_{ij} $(j \neq i)$ are chosen to be positive, the matrix is irreducible.

In total for an N-state Markov chain this comes to $(N^2 + 3N - 2)/2$ parameters. By choosing to work with reversible chains, we have reduced the number of parameters considerably.

- *Method:* a day-by-day one country fit with a fully data imposed 11-state Markov structure.
- *Data used:* GBP only.
- *Markov structure:* data imposed structure, symmetrisable matrices only.
- *Parameters in the minimisation vector θ:*

  g vector, α, above diagonal elements of A and invariant mass vector π

6.3.5 Results

The extensive tests we have carried out indicate that the best fits to market data were obtained using an entirely data imposed underlying structure. Therefore, to begin with we present a summary table (Table 6.2) showing the fits obtained for all the data periods discussed in §6.2, using a day-by-day calibration with an entirely data imposed underlying structure. The results presented in this table were obtained using a 15-state Markov chain, with all maturities weighted equally and the covariance matrix, V, taken to be the identity.

The table gives the mean and standard deviation of the fits for each 100 day period. It also reports the minimum, median, maximum and first and third quartiles of the daily fits during the periods. All values are quoted in terms of basis points. For reference we also indicate how many base rate changes occur in each period.

Perhaps the most interesting column in this table is the *Median* column which presents the median values of the sum of absolute errors for each day's fit. This sum consists of 8 terms, one for each maturity, so the bp error per maturity is 1/8 of the figure given. The results in the table suggest that 1992 is a difficult year to fit; the mean and standard deviations are high and the median error per maturity is 2 bp whereas for most of the other periods it is 1 bp or less. The poor fit for 1992 is purely down to the freak events that occurred in the UK economy in this year and no matter which (realistic mathematical) model you choose, nothing will cope well under these circumstances. Generally speaking the results are very good, even looking at the upper quartile we find that only in three of the 19 periods did the error exceed 2 bp per maturity.

	Calendar Period	Day numbers	BRC	Statistics of day-by-day calibration, all values in bp						
				Mean	Std. Dev.	Min	Q1	Median	Q3	Max
1	17th Feb 1992 - 15th July 1992	30-129	1	18.538	6.453	10.076	13.796	16.565	21.385	35.305
2	5th May 1992 - 25th Sept 1992	80-179	2	19.015	8.334	9.395	13.768	16.543	21.672	69.703
3	16th July 1992 - 7th Dec 1992	130-229	3	18.196	26.629	0.001	6.939	14.672	21.268	226.958
4	28th Sept 1992 - 18th Feb 1993	180-279	3	10.238	6.147	1.70	5.773	9.44	13.117	26.423
5	8th Dec 1992 - 7th May 1993	230-329	1	9.677	4.77	0.975	6.107	9.443	12.852	21.472
6	19th Feb 1993- 19th July 1993	280-379	0	6.678	4.613	0.001	2.94	6.422	8.992	23.569
7	10th May 1993 - 28th Sept 1993	330-429	0	5.261	3.382	0.128	2.874	4.5	7.631	15.704
8	20th July 1993 - 8th Dec 1993	380-479	1	5.397	3.488	0.055	2.692	4.463	8.144	15.510
9	29th Sept 1993 - 23rd Feb 1994	430-529	2	6.310	3.949	0.12	3.358	5.181	9.322	16.944
10	9th Dec 1993 - 11th May 1994	480-579	1	7.248	4.738	0.039	4.099	6.833	9.466	31.231
11	24th Feb 1994 - 22 July 1994	530-629	0	9.649	6.118	0.164	5.39	8.263	13.457	30.662
12	12th May 1994 - 4th Oct 1994	580-679	1	10.005	5.32	0.443	6.337	9.482	13.508	27.236
13	25th July 1994 - 16th Dec 1994	630-729	2	7.02	4.188	0.558	3.646	6.328	10.086	18.910
14	5th Oct 1994 - 6th Mar 1995	680-779	2	8.402	4.03	1.036	5.273	8.463	11.09	21.629
15	19th Dec 1994 - 23rd May 1995	730-829	1	8.437	2.608	0.77	6.666	8.243	10.081	15.573
16	7th Mar 1995 - 3rd Aug 1995	780-879	0	7.846	3.370	0.795	5.78	8.132	9.687	17.092
17	24th May 1995 - 16th Oct 1995	830-929	0	4.524	3.286	0.081	2.152	3.733	5.952	16.660
18	4th Aug 1995 - 28th Dec 1995	880-979	1	2.586	1.844	0.001	1.058	2.231	3.576	8.969
19	17th Oct 1995 - 8th Mar 1996	930-1029	3	5.503	3.436	0.001	2.387	4.956	8.552	14.726

Table 6.2: Results of fits for the day-by-day calibration using the 19 sample periods of 100 days discussed in §6.2. Note that these results were obtained using a 15-state data imposed underlying Markov chain. The column marked 'mean' above refers to the average bp error between model and observed values per day. The standard deviation is that of the daily bp error in the period. Q1 and Q3 denote the first and third quartiles. Min, Median, Max again refer to the bp error per day. BRC is used to denote the number of Bank of England base rate changes.

As mentioned in §6.2 for more detailed analysis we chose to focus on period 14, 5th October 1994 to 6th March 1995. This period does not have the freak events present in 1992 but still contains 2 base rate changes and so we can examine how these affect our model's fit. In Table 6.3 we present summary statistics for the cases A, B and C. Figures 6-7, 6-8 and 6-9 show the diagnostics from the runs.

Diagnostic plots for day-by-day calibration case A: OU structure

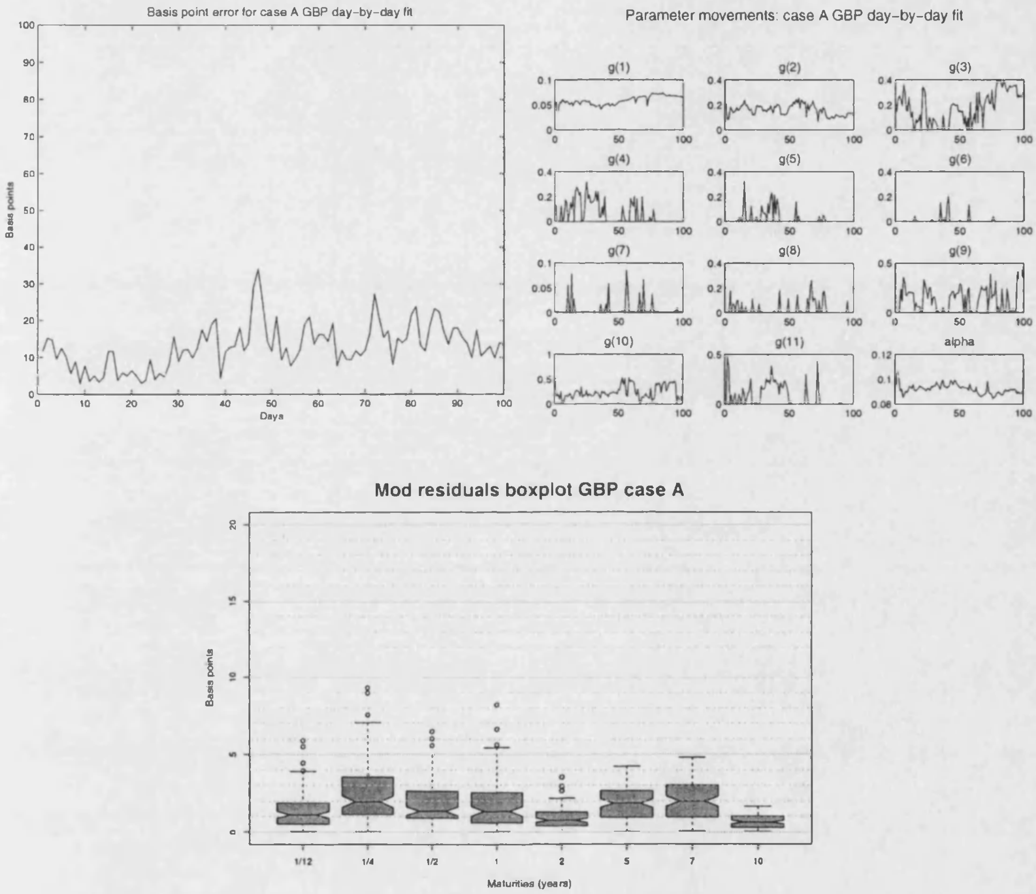

Figure 6-7: Diagnostic plots for the day-by-day calibration for case A, which is described in §6.3.2. The bp error plot (top left) shows the total error, given in bp, between the market and model yield curves for each fitted day. The evolution of the parameters g and α over the whole fitting period are given in the top right plot. Finally we give boxplots showing the mean and quartiles of the mod residuals for each maturity.

Diagnostic plots for day-by-day calibration case B: data imposed OU structure

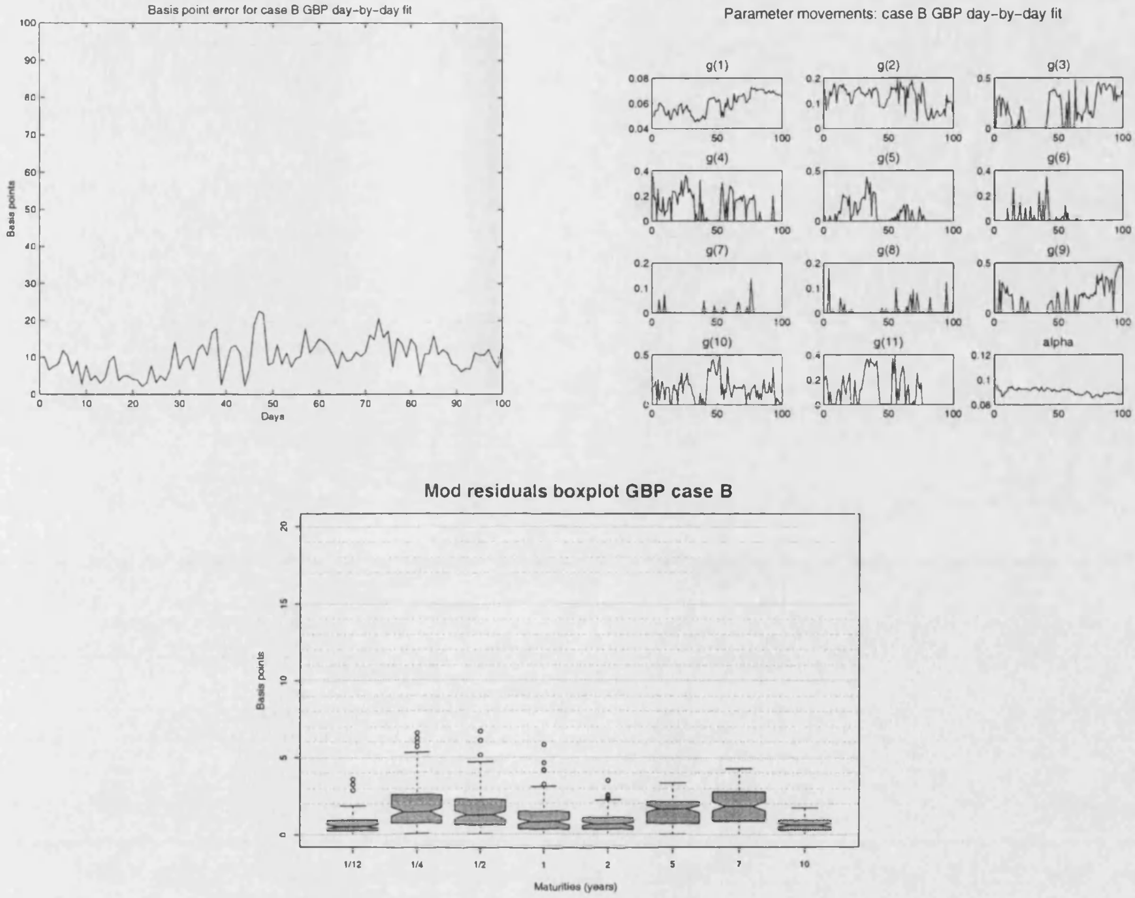

Figure 6-8: Diagnostic plots for the day-by-day calibration for case B, which is described in §6.3.3. The bp error plot (top left) shows the total error, given in bp, between the market and model yield curves for each fitted day. The evolution of the parameters **g** and α over the whole fitting period are given in the top right plot. Finally we give boxplots showing the mean and quartiles of the mod residuals for each maturity.

Diagnostic plots for day-by-day calibration case C: data imposed structure

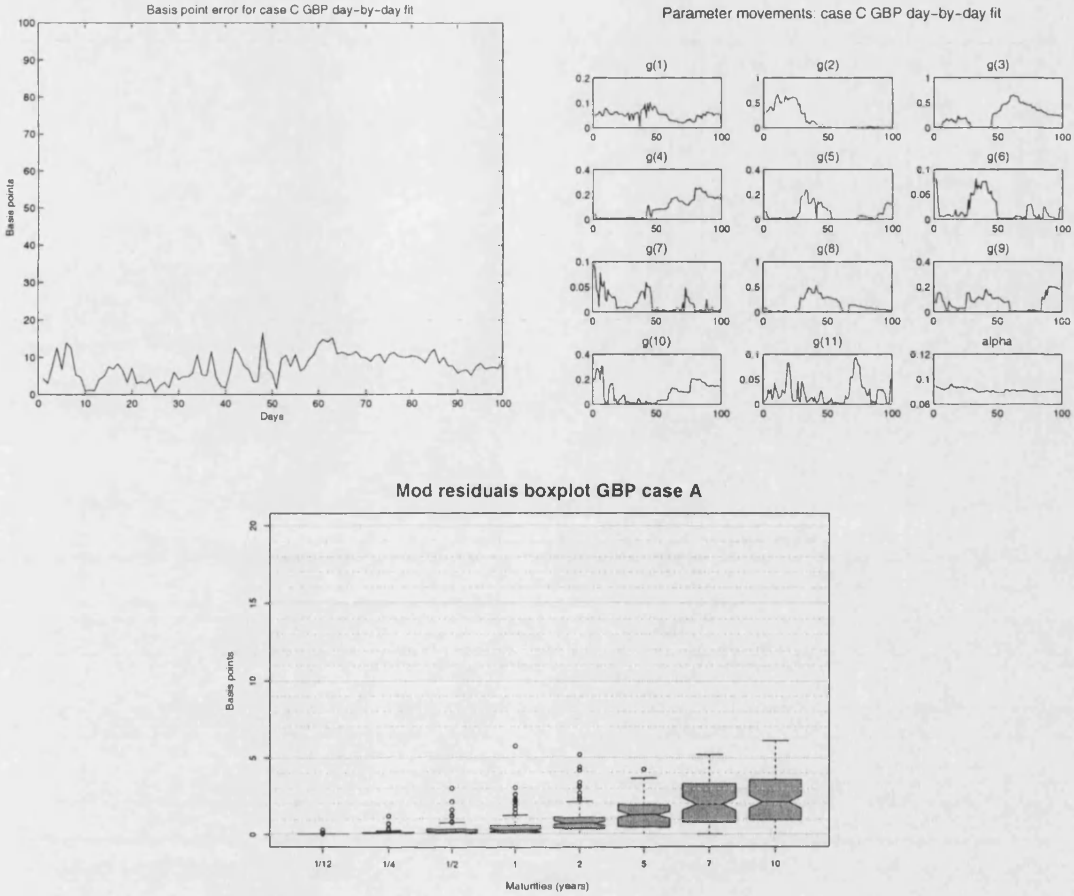

Figure 6-9: Diagnostic plots for the day-by-day calibration for case C, which is described in §6.3.4. The bp error plot (top left) shows the total error, given in bp, between the market and model yield curves for each fitted day. The evolution of the parameters g and α over the whole fitting period are given in the top right plot. Finally we give boxplots showing the mean and quartiles of the mod residuals for each maturity.

As we claimed at the beginning of the results section, we see from the basis point error

Day-by-day calibration statistics GBP (all values are in basis points)							
	Mean	Std. Dev.	Min	Q1	Median	Q3	Max
Case A	12.855	6.033	2.816	9.319	12.301	16.178	34.075
Case B	9.939	4.460	2.143	6.853	10.205	12.640	22.577
Case C	7.585	3.588	0.693	4.842	7.302	10.400	16.619

Table 6.3: This table contains summary statistics relating to the three cases explored using the day-by-day calibration procedure, with an 11-state Markov chain using GBP data.

plots in Figures 6-7, 6-8 and 6-9 that the more we allow the data to influence the underlying structure, the better the quality of the fit.

The parameter movement plots for each of the cases show the evolution of the parameter α over the fitting period and also plots the evolution of the normalised g vector. For g, we plot

$$\frac{g(i)}{\sum_{j=1}^{N} g(j)} \tag{6.12}$$

for $i = 1, .., N$, where N is the number of states in the Markov chain. The parameter values appear to fluctuate a little over the course of the fitting period and this lack of stability is something we would wish to reduce in our subsequent calibration exercises.

Finally, the box plots indicate that the median fit for each maturity is consistently at or below the 2 bp mark.

The boxplots also indicate that bonds with some maturities are harder to fit than others. In this situation, we could alter the weight on the difficult maturities by modifying the covariance matrix so that the minimisation would focus more on improving the fit for a particular maturity.

6.3.6 Discussion

The modelling exercises in this section have been concerned with the type of underlying structure we should use for the Markov chain potential models. The efforts have concentrated on data imposed structures and the motivation for this is twofold. Firstly, there does not appear to be any other structure which has compelling reasons for being chosen. Secondly, using a fully data imposed structure should enable the model to understand and adapt to the market more easily.

All the underlying structures considered have provided impressive results. In particular for case C, over a 100 day fitting period with 8 ZCB prices of different maturities, we can obtain a mean basis point error per derivative of around 1 bp. A model that is fitting yields to within a basis point is good enough to trade off, and we are here getting close to that degree of precision with so few states in our chain. This is obviously an encouraging starting point and our next objective must be to perform a calibration using more realistic assumptions.

6.4 Modelling Exercise 2: rigid calibration with out-of-sample tests

In this approach, which we refer to as a **rigid calibration**, we consider calibrating the model to a period of K days' data using an approximation of the likelihood (6.2). This calibration procedure is more "honest" than the day-by-day calibration, in that it requires the parameters to be the same over all days and this is more consistent with the theoretical assumptions underlying these models. The simplification used here involves replacing (6.3) (which involves a sum over all possible paths of the chain during the K days of the calibration period) by a single term corresponding to a path which remains in its initial state throughout the calibration period. The true calibration, involving a sum over all possible paths of the underlying chain during the K days, would computationally be far too slow.

The assumption that the initial state does not change is reasonable provided the length of the calibration period is no more than a few days and the period is one in which a change in the underlying state is unlikely. If the calibration period has a number of base rate changes and is generally considered a period of high volatility, then this assumption is less acceptable and frequent recalibration would be necessary.

Taking the initial state to be x, the likelihood function in this case will be

$$\begin{aligned}\Lambda_K(\mathbf{X}_K, \mathbf{y}_K, \theta) &= \frac{\pi(x) f_0(\theta)}{(2\pi|V|)^{K/2}} \prod_{j=1}^{K} p_{xx}(s_j;\theta) \exp[-b(y_j - Y(x;\theta))] \\ &\simeq \frac{\pi(x) f_0(\theta)}{(2\pi|V|)^{K/2}} \prod_{j=1}^{K} \exp\left(-q_{xx}s_j\right) \exp\left[-b(y_j - Y(x;\theta))\right] \\ &= \frac{\pi(x) f_0(\theta)}{(2\pi|V|)^{K/2}} \exp\left[-\sum_{j=1}^{K} \{b(y_j - Y(x;\theta)) + q_{xx}s_j\}\right],\end{aligned} \tag{6.13}$$

where $-q_{xx}$ is the diagonal entry in the x position of the Q-matrix Q. Therefore, we can now find the θ which maximises the $\log \Lambda_K(\mathbf{X}_K, \mathbf{y}_K, \theta)$; this value is denoted θ^*.

Again since the particular initial state is not important, we may as well assume that it is the first one labelled. Hence, in this case we obtain θ^* by simply minimising

$$\sum_{j=1}^{K}[b(y_j - Y(1;\theta)) + q_1 s_j] - \log f_0(\theta), \tag{6.14}$$

where $-q_1$ is the diagonal entry in the first position of the Q-matrix Q.

Having found our calibration vector θ^*, we can now test the model out-of-sample by taking the days after the calibration period and trying to fit the yield curves by only allowing changes in the posterior distribution of X. We compute this distribution, $\pi_n(X_n, \mathbf{y}_n)$, for any day n after the calibration period using the recursive formulae

$$\pi_n(x, \mathbf{y}_n) \propto \sum_{\xi} \pi_{n-1}(\xi, \mathbf{y}_{n-1}) p_{\xi x}(s_n; \theta_n^*) \exp[-b(y_n - Y(x; \theta_n^*))], \tag{6.15}$$

for all $X_n \in \mathcal{X}$. Observe that we use the proportional sign because the above is subject to normalisation; we insist $\sum_{i=1}^{N} \pi_n(X_n(i), \mathbf{y}_n) = 1$ for all n. The initial distribution, π_0 is a point mass on the state chosen as the initial state in the calibration exercise[3].

This approach to calibration may be looked at as quasi-Bayesian and uses estimates for the posterior distribution π_n of the underlying Markov chain X_n at time n. Therefore, to price a derivative on day n, we shall use the expression

$$\sum_{x} \pi_n(x, \mathbf{y}_n) F(x, \theta_n^*), \tag{6.16}$$

where $F(x, \theta)$ is the price which the Markov chain potential model would produce if the starting state were x and the true parameter value were θ. This would apply, for example, to the pricing of zero-coupon bonds; so, in particular, we end up with a continuum of possible yield curves at any given time, even though the model with known θ could only produce one yield curve for each possible state of the Markov chain.

6.4.1 Case A: initial calibration only

Here we consider the K days of data which precede the first day of our focus data period (period 14) and we use the method described above to calibrate the model. We will only consider entirely data imposed underlying structures so the calibration procedure can impose its own Q-matrix (see §6.3.4). Once the initial-period calibration is

[3]The reader should refer to §6.5 for a more complete justification of (6.15)

complete and we have found, θ^*, our 'optimum estimates' for the model parameters θ, we will then hold this fixed and carry out an out-of-sample test by investigating the fit of yield curves for the 100 days in period 14.

- *Method:* a rigid calibration of the yield curve for one country with a data imposed 11-state Markov structure.
- *Markov structure:* data imposed structure.
- *Data used:* GBP only.
- *Parameters in the minimisation vector θ:*

  g vector, α vector, upper triangular half of A, invariant mass π.

- *Other settings:*

Parameter	Value
Calibration period, K	5 days (670-679)
Total fitting period	100 days (680-779)

6.4.2 Case B: 10-day periodic re-calibration

In this case we investigate a periodic re-calibration of our rigid calibration procedure. Practitioners are likely to have reservations over the philosophy in case A that a model's parameters remain constant over such a large period. It could perhaps be argued that exogenous influences force the market to change and so the parameters of any model should be allowed to change and adapt in this situation.

One adaptation of case A is the following. Calibrate the model over K days worth of data allowing the underlying structure to be entirely data imposed (see §6.3.4) then fit the model to J days of data by only allowing the posterior distribution to vary using the recursive prescription (6.15).

Next, having fitted for J days we repeat the procedure. So recalibrate the model over K days prior to the day we are now at and then (again) fix and use the θ^* obtained in the calibration to fit over the next J days. This periodic re-calibration is repeated continuously until the total number of days we have fitted over is 100.

- *Method:* a rigid calibration approach with periodic recalibration, one country fits using an 11-state Markov chain potential model.
- *Markov structure:* data imposed structure.
- *Data used:* GBP only.
- *Parameters in the minimisation vector θ:*

  g vector, α, upper triangular part of A, invariant mass π.

- *Other Settings:*

Parameter	Value
Calibration period, K	5 days
Re-calibrate after fitting, J	10 days
Total Fitting Period	100 days

6.4.3 Case C: 1-day out-of-sample fits

A special case of the periodic re-calibration example give in case B is to re-calibrate the model after fitting each day. This then becomes a standard one-day out-of-sample fitting exercise using our rigid calibration approach.

- *Method:* a rigid calibration approach, recalibrating after each day's fit, on one country using an 11-state Markov chain potential model.
- *Markov structure:* data imposed structure.
- *Data used:* GBP only.
- *Parameters in the minimisation vector θ:*

  g vector, α, upper triangular part of A, invariant mass π.

- *Other Settings:*

Parameter	Value
Calibration period, K	5 days
Re-calibrate after fitting, J	1 day
Total Fitting Period	100 days

6.4.4 Results

The results of this calibration exercise are quite poor when compared to those of the day-by-day fit. This is not surprising when you consider that we have essentially asked the model to fit just five days worth of data and then to run with that unaltered for the next 100 days. The industry would never contemplate doing what we have attempted in case A and only calibrate their models every 4-5 months (100 days); the results show why.

Table 6.4 gives summary statistics for this calibration method, recalibrating every J days using $K = 5$ days of data. This table indicates that the median error in bp per maturity is of the order of 35 for $J = 100$ and of the order of 6 for $J = 1$, the first is useless, the second is poor. We give more detailed diagnostics in Figures 6-10 and 6-11, corresponding to the cases $J = 100$, $J = 10$ and $J = 1$.

Figure 6-10 displays behaviour we would expect. For example, the basis point error plot for case A ($J = 100$) shows a gradual deterioration in the quality of the fit over the 100 day fitting period. The results of case B ($J = 10$) are much improved and the corresponding basis point error plot for this case shows that the quality of the fit worsens over the 10 day fitting period before being pulled back down by the recalibration. We therefore see cycles of peaks and troughs. This behaviour, although expected, is undesirable especially given the extent of the variations in the quality of the fit.

Rigid calibration statistics, GBP (all values are in basis points) 5 day calibration period							
Re-calibrate After	Mean	Std. Dev.	Min	Q1	Median	Q3	Max
100 days	257.09	101.87	42.24	175.79	270.46	349.67	416.87
50 days	140.04	68.08	34.50	90.40	124.44	180.32	313.13
25 days	105.78	54.04	21.64	63.58	95.44	136.31	246.78
10 days	94.12	53.45	19.73	49.82	82.93	128.65	232.33
5 days	76.17	45.65	19.71	41.97	63.46	99.50	200.14
2 days	60.11	38.63	7.34	30.51	48.62	74.47	192.79
1 day	55.34	33.99	7.22	29.46	47.81	66.60	162.49

Table 6.4: This table contains summary statistics for the rigid calibration procedure. The results are for fits over a 100 day period using different re-calibration intervals. All results are for an 11-state chain using GBP data.

Daily basis point error plots for the rigid calibration

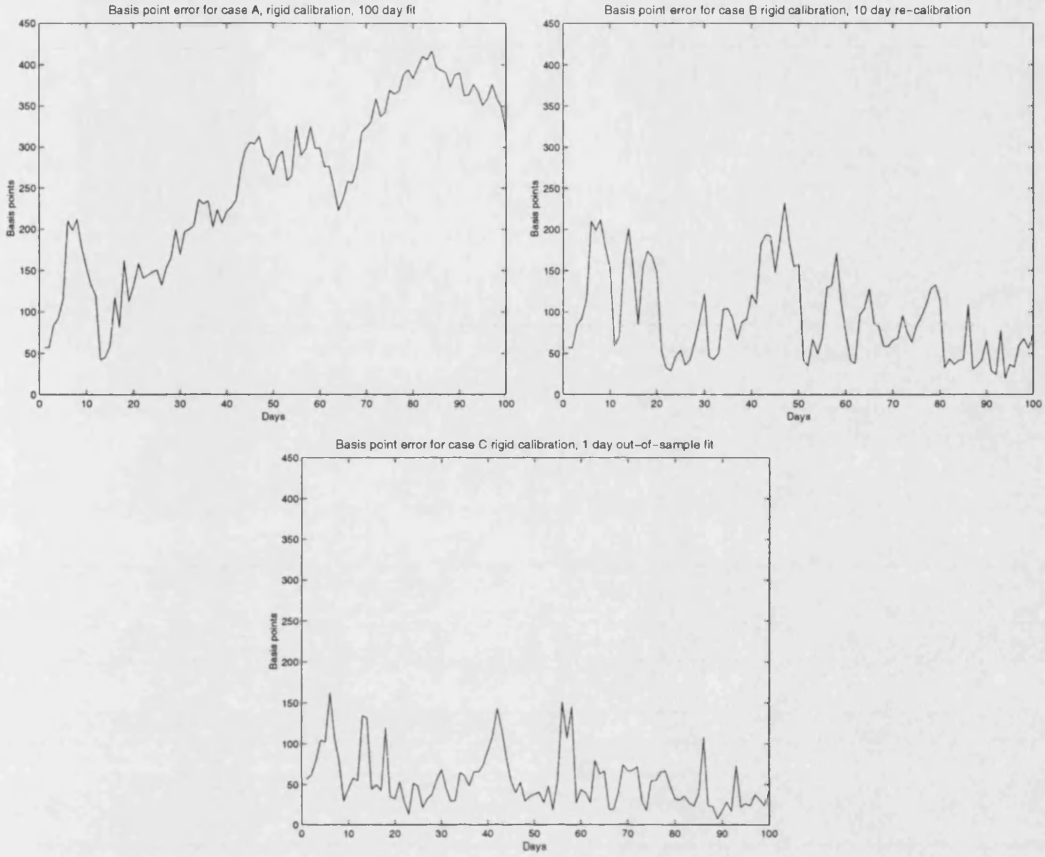

Figure 6-10: We show three basis point error plots referring to cases A, B and C. These plots show the cumulative error in basis points for each day over the 100 day fitting period.

Notice that the boxplots in figure 6-11 indicate that case C, the 1-day out-of-sample scenario is the only one in which the basis point error is consistent across the maturities.

Boxplots of mod residuals for the rigid calibration

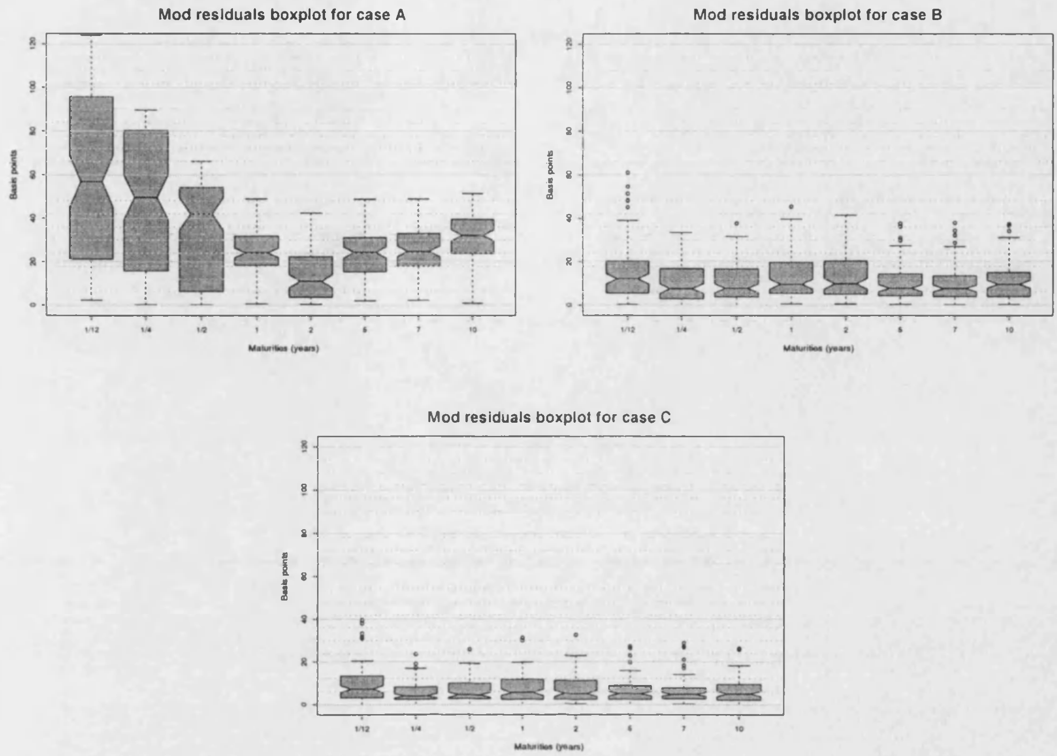

Figure 6-11: We give three boxplots showing the characteristics of the mod residuals for each maturity for the cases A, B and C.

6.4.5 Discussion

The most acceptable of the trials suggested here is the 1-day out-of-sample fit considered in case C. This is the nearest to the accepted practice in the industry as it amounts to re-calibrating interest rate models each night. The industry and practitioners find it much more reassuring if a model's 'knowledge' of the market is updated every day. The quality of the fit in this example is the best of the three cases, in fact, the mean basis point error for case C is five times better than that of case A. We also see more satisfactory boxplots of the mod residuals in Figure 6-11.

Generally speaking, we would prefer to see better results than we have achieved in these trials. There are a number of changes we could make to try and improve the fit. For example, we could try to calibrate the model using specific days rather than just a period of consecutive days before the fitting period. One option is to choose the 5 Mondays prior to the first fitting day to calibrate the model. Intuitively, we would hope that by calibrating the interest rate model to a longer time interval of data, the model parameters would be closer to their "correct" values. However, if we adopted this approach, we would also want to have a different start state for each Monday. This is not computationally feasible as we could not sum over all paths.

The assumptions we have used here are still not as convincing as the scenario laid out in §6.1.2. Moreover, the restriction in the calibration that the path remains in its initial state for the whole of the calibration period is not ideal. Obviously, an extension of the calibration would be to allow the chain just one jump during the K days, but instead of pursuing this line of thought, it is better to investigate whether we could do better than this approach altogether.

6.5 Modelling Exercise 3: conditional-independence (CI) and random walk (RW) calibration

The third approach to calibration we shall consider is the closest to the methodology outlined in §6.1.2. The motivation we use here is that having seen a large amount of data, we should have a fairly good idea what the values of the parameters must be; the values of θ will largely be determined by the long-run historical average behaviour of the system. Conversely, we imagine the posterior distribution of X_n will be more influenced by the recent history of the Markov chain. Hence we use an approximate conditional independence where the recent history tells us all we know of X_n and distant history tells us about θ. Therefore we imagine the situation where there has been a large amount of observed data and we postulate that

$$L_n(x, \mathbf{y}_n, \theta) = \pi_n(x, \mathbf{y}_n) l_n(\theta, \mathbf{y}_n). \tag{6.17}$$

We further assume that

$$l_n(\theta, \mathbf{y}_n) \propto \exp(-\frac{1}{2}(\theta - \hat{\theta}_n)^T S_n^{-1}(\theta - \hat{\theta}_n)) \tag{6.18}$$

for some positive-definite symmetric matrix S_n^{-1}. We may find the value of S_n^{-1} by computing the second derivative matrix of l_n with respect to θ. However, for simplicity we only consider the diagonal matrix S_n^{-1} and ignore cross derivative terms. The quadratic approximation to the likelihood (6.18) is quite natural if we think we have nearly identified the true value of θ.

The values $\hat{\theta}_n$, S_n^{-1}, and $\pi_n(\cdot, \mathrm{y}_n)$ are computed recursively, using the assumed form (6.17) of the likelihood. Supposing that we already know $\hat{\theta}_{n-1}$, S_{n-1}^{-1}, and $\pi_{n-1}(\cdot, \mathrm{y}_{n-1})$, then (6.4) and (6.17) give us that

$$\begin{aligned} L_n(x, \mathrm{y}_n, \theta) &= \sum_\xi \pi_{n-1}(\xi, \mathrm{y}_{n-1}) l_{n-1}(\theta, \mathrm{y}_{n-1}) p_{\xi x}(s_n; \theta) \exp[-b(y_n - Y(x;\theta))] \\ &\propto \sum_\xi \pi_{n-1}(\xi, \mathrm{y}_{n-1}) p_{\xi,x}(s_n; \theta) \exp[-b(y_n - Y(x;\theta))] \\ &\qquad . \exp[-\frac{1}{2}(\theta - \hat{\theta}_{n-1}) \cdot S_{n-1}^{-1}(\theta - \hat{\theta}_{n-1})] \end{aligned} \tag{6.19}$$

we sum this expression over x, and numerically pick θ to maximise it; the maximising value is our new estimate $\hat{\theta}_n$ of θ. As in the previous calibration we compute $\pi_n(X_n, \mathrm{y}_n)$, for any day n using the recursive formulae

$$\pi_n(x, \mathrm{y}_n) \propto \sum_\xi \pi_{n-1}(\xi, \mathrm{y}_{n-1}) p_{\xi x}(s_n; \hat{\theta}_n) \exp[-b(y_n - Y(x; \hat{\theta}_n))], \tag{6.20}$$

We use this approximation to the likelihood because the correct expression would involve integrating (6.19) with respect to θ which computationally is not feasible. Instead, to avoid a multi-dimensional integral we assume that the posterior distribution for θ can be replaced by the point mass at $\hat{\theta}_n$. The above method will be referred to as a conditional-independence (CI) calibration.

6.5.1 Case A: conditional-independence calibration

This experiment will calculate the MLE (maximum likelihood estimate) using the prescription in (6.19). The initial value S_0^{-1} is taken to be the identity matrix, the vector $\hat{\theta}_0$ is taken as a zero vector and the initial posterior distribution is the uniform distribution, so $\pi_0(x) = 1/N$ for all $x \in \mathcal{X}$, where $N = |\mathcal{X}|$. Once we have calculated our point estimates, $\hat{\theta}$, we use the recursive formula (6.20) to calculate the posterior distribution. Again for this quasi-Bayesian approach, a derivative on any day n will be priced using

$$\sum_x \pi_n(x, \mathrm{y}_n) F(x, \hat{\theta}_n), \tag{6.21}$$

where $F(x, \theta)$ is the price which the Markov chain potential model would produce if the starting state were x and the true parameter value were θ.

- *Method:* a one country CI calibration using an 11-state Markov chain.
- *Markov structure:* data imposed structure.
- *Data used:* GBP only.
- *Parameters in the minimisation vector θ:*

g vector, α vector, upper triangular part of A and π vector.

6.5.2 Case B: random walk calibration

The random walk (RW) calibration is very similar to the CI method which was described above and it can be seen that the CI method is a special case of the RW calibration. The idea here is that we would now like our model to take account of changes which may occur in the market from day to day. To do this, we allow the vector θ to change according to a Gaussian random walk with zero mean and variance matrix R. The method we use is similar to the Kalman filter argument.

For ease of notation let Y be the discrete time vector process representing the yields generated by our model and θ_n the parameter vector at time n. Consider modelling the evolution of these processes according to the linear model

$$\begin{aligned} \theta_n &= \theta_{n-1} + \varepsilon_n \qquad (6.22) \\ Y_n &= C\theta_n + \eta_n, \qquad (6.23) \end{aligned}$$

where for every n, $\varepsilon_n \sim N(0, R)$ and $\eta_n \sim N(0, V)$ are independent random variables and C is a matrix. If $\mathcal{Y}_n$ denotes the σ-field generated by $\{Y_k : k \leq n\}$, and if we have that conditional on $\mathcal{Y}_n$ the law of θ_n is $N(\hat{\theta}_n, S_n)$, then

$$\left(\begin{matrix} \theta_{n+1} \\ Y_{n+1} \end{matrix} \middle| \mathcal{Y}_n \right) \sim N\left(\begin{pmatrix} \hat{\theta}_n \\ C\hat{\theta}_n \end{pmatrix}, \begin{pmatrix} R + S_n & (R + S_n)C^T \\ C(R + S_n) & V + C(R + S_n)C^T \end{pmatrix} \right), \qquad (6.24)$$

and likewise

$$\left(\begin{matrix} \theta_{n+1} \\ Y_{n+1} - C\theta_{n+1} \end{matrix} \middle| \mathcal{Y}_n \right) \sim N\left(\begin{pmatrix} \hat{\theta}_n \\ 0 \end{pmatrix}, \begin{pmatrix} R + S_n & 0 \\ 0 & V \end{pmatrix} \right). \qquad (6.25)$$

It is a straightforward exercise to confirm from (6.24) that the law of θ_{n+1} given $\mathcal{Y}_{n+1}$ is $N(\hat{\theta}_{n+1}, S_{n+1})$, where $\hat{\theta}_{n+1}$ is the value of θ maximising the joint density of the distribution (6.24) or equivalently (6.25):

$$\exp\left[-\frac{1}{2}(\theta - \hat{\theta}_n)^T (R + S_n)^{-1}(\theta - \hat{\theta}_n) - \frac{1}{2}(y - C\theta)^T V^{-1}(y - C\theta) \right], \qquad (6.26)$$

and $-S_{n+1}^{-1}$ is the second derivative of the log-likelihood with respect to θ. The actual estimation problem we face has non-linear dynamics, but we shall suppose that a local linear approximation is adequate, so we replace (6.23) with

$$Y_n = Y(X_n, \theta_n) + \eta_n,$$

giving the analogue of (6.26) to be

$$\exp\left[-\frac{1}{2}(\theta - \hat{\theta}_n)^T (R + S_n)^{-1}(\theta - \hat{\theta}_n) - \frac{1}{2}(y - Y(X_n, \theta))^T V^{-1}(y - Y(X_n, \theta)) \right] \qquad (6.27)$$

Taking $R = 0$ corresponds to the CI calibration discussed previously and this is equivalent to dropping the random walk contribution in (6.22), we find that (6.27) reduces to the exponential terms in (6.19).

For the RW calibration, we shall take $R = (\beta^{-1} - 1)S_n$, for a fixed variable $\beta \in (0, 1)$, the likelihood (6.19) now becomes

$$\sum_{\xi} \pi_{n-1}(\xi, \mathrm{y}_{n-1}) p_{\xi,x}(s_n; \theta) \exp[-b(y_n - Y(x; \theta))]$$
$$.\exp[-\frac{\beta}{2}(\theta - \hat{\theta}_{n-1}) \cdot S_{n-1}^{-1}(\theta - \hat{\theta}_{n-1}))]. \qquad (6.28)$$

The introduction of the parameter β allows our model to take account of changes to the market environment and at the same time provides us with a mechanism for controlling the level in which new observations are allowed to influence the model. One of the other advantages of this approach is that the implementation of it is virtually identical to that of the CI calibration.

- *Method:* a RW calibration using an 11-state Markov chain.
- *Markov structure:* data imposed structure.
- *Data used:* GBP only.
- *Parameters in the minimisation vector θ:*

g vector, α vector, upper triangular part of A and π vector.

6.5.3 Results

The results obtained for the fits investigated in this section were very good bearing in mind we have only considered an 11-state underlying Markov chain. For the CI calibration on a one country (GBP) eight point yield curve we obtain a median error of just over 5 bp per maturity which is already better than even the 1-day rigid fit. We also observe (from Table 6.5) that the best daily fit we achieve has an error of approximately only 1 bp per maturity. On the other hand, the worst day fitted in the 100 day period was on average around 13 bp out per maturity, which is considerably less appealing.

Diagnostics of the CI fit are given in Figure 6-12. Notice that in the basis point error plot, days 35 to 55 are much harder to fit than the other days in the interval. This period contains one of the two base rate changes[4] that occur over the 100 days. After day 55 the fit improves marginally before deteriorating again to coincide with the second base rate change. This observation together with the other characteristics of the fit lead us to conclude that the CI calibration performs less effectively under the volatile circumstances leading to a base rate change. If this is indeed the case, an obvious cause may be that the CI calibration assumes that our knowledge of the parameter vector θ is being updated over time, not that θ is necessarily changing or evolving in time. Indeed, the plot of parameter movements in this calibration exercise reveal that the parameters are quite tightly constrained.

CI/RW calibration statistics one country GBP (all values are in basis points)							
β	Mean	Std. Dev.	Min	Q1	Median	Q3	Max
1.0 (CI)	47.936	22.253	8.441	30.155	41.755	65.786	107.383
0.8	34.995	14.942	8.432	23.944	34.272	45.493	78.581
0.6	26.398	11.243	7.734	18.004	25.72	34.134	62.514
0.4	23.957	10.172	5.751	15.95	23.216	30.261	53.292
0.2	21.346	9.48	5.462	13.35	20.529	26.704	45.744
0.1	20.339	9.31	4.917	13.062	18.942	26.088	45.509

Table 6.5: This table contains summary statistics relating to the one country (GBP) period 14 CI/RW fits for different β values. Note that the case $\beta = 1.0$ corresponds to the CI calibration.

[4]base rate changes occur on 7th December 1994 (day 43) and 2nd February 1995 (day 79).

Diagnostic plots for case A: GBP CI calibration

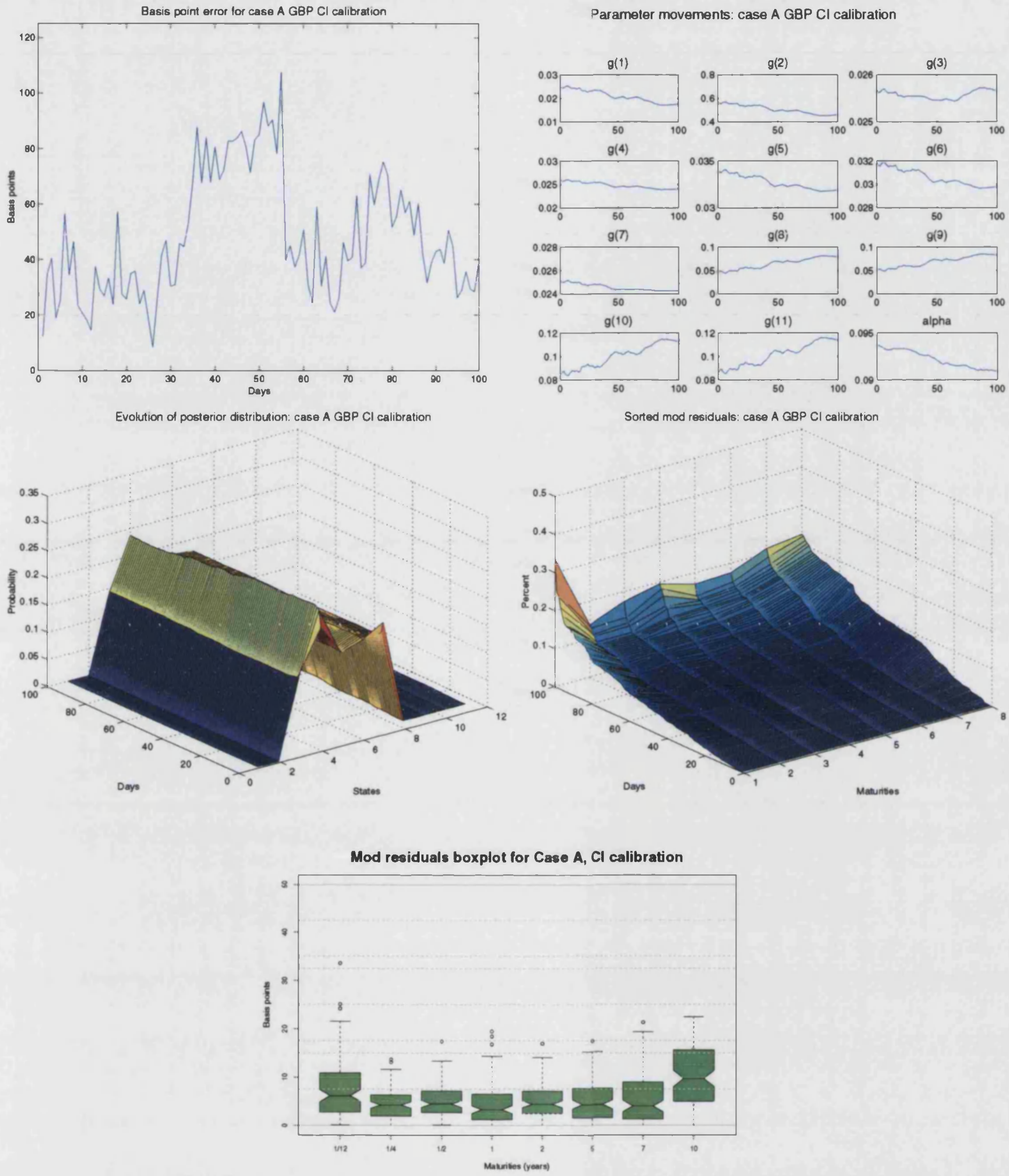

Figure 6-12: These plots relate to the one country CI calibration described in case A. The basis point error plot shows the cumulative error in basis points for each day over the 100 day fitting period. The parameter change plot shows how the $\mathbf{g}$ vector and α scalar change over the 100 day fitting period. We give a surface plot which shows the evolution of the posterior distribution over the fit. The characteristics of the mod-residuals are given in the boxplots.

Diagnostic plots for case B: GBP RW calibration ($\beta = 0.2$)

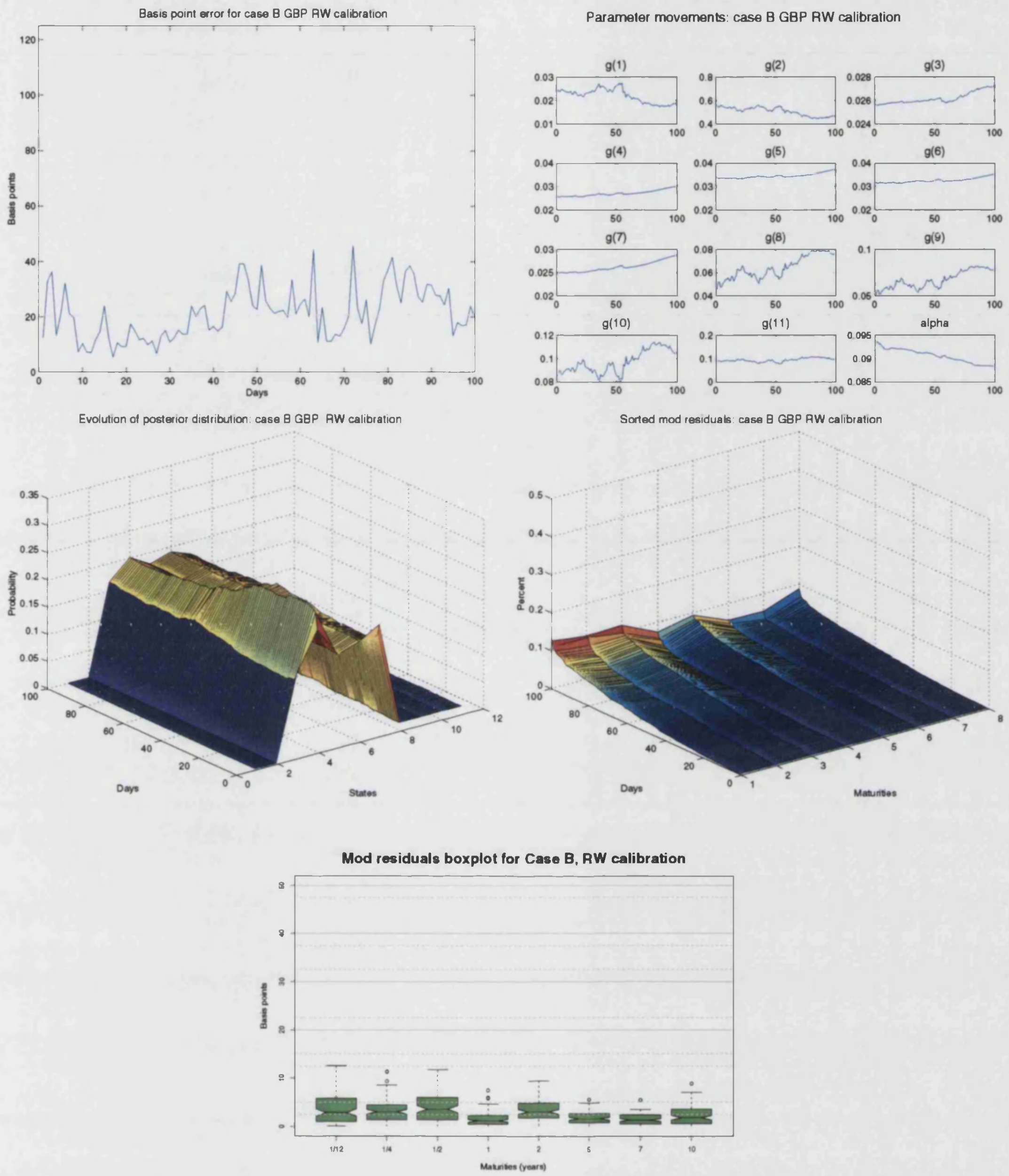

Figure 6-13: These plots relate to the one country RW calibration described in case B with $\beta = 0.2$. The basis point error plot shows the cumulative error in basis points for each day over the 100 day fitting period. The parameter change plot shows how the $\mathbf{g}$ vector and α scalar change over the 100 day fitting period. We give a surface plot which shows the evolution of the posterior distribution over the fit. The characteristics of the mod-residuals are given in the boxplots.

The plots showing the evolution of the posterior distribution do not reveal a great deal.

However, one observation we can make is that the states seems to fall into two classes. There is a set of states which have little or no initial weight attached to them and another set for which all the states broadly speaking have an even spread of mass.

We do not present any analysis of the underlying structure which was chosen by the data and whose values are altered by changing elements of θ. The reason for this is that there is little that can be said about the structure other than the properties that we have insisted it satisfy. Something which is considerably more interesting is the breakdown in the quality of the fit per maturity. We provide two plots to examine this. The surface plot is of the sorted mod-residuals. The boxplot also shows the characteristics of the mod residuals and we see from this that with the exception of maturity 8 (the ten year bond) the mean fit for each of the maturities is around the same mark (5 bp). There are also one or two clear outliers.

Unsurprisingly we observe from the statistics in Table 6.5 that the fit improves considerably as we take lower values of β in the term

$$\exp[-\frac{\beta}{2}(\theta - \hat{\theta}_{n-1}) \cdot S_{n-1}^{-1}(\theta - \hat{\theta}_{n-1}))] \tag{6.29}$$

contained within (6.28). Performing an RW calibration with $\beta = 0.2$, we can obtain median errors of around 2.5 bp per maturity which is significantly better than the fit obtained from the CI calibration.

Figure 6-13 contains plots for the RW calibration where $\beta = 0.2$. From the plot of the basis point error, we see immediately that the RW calibration method provides a consistently better fit over the 100-day interval than we saw from the CI fit. Next observe that when we compare the plot showing parameter movements in $\mathbf{g}$ and α, with the corresponding plot for the CI case given in Figure 6-12, we find that even with this seemingly low value of β we obtain quite impressive stability. The boxplot of mod residuals for the RW case confirm the improved median fit for each of the maturities, though the quality of fit between maturities is slightly less consistent.

6.5.4 Discussion

The conditional-independence and random walk calibrations introduced in this modelling exercise seem to be as close as we can get to the ideal methodology explained in §6.1.2. If we were to attempt a more realistic and ambitious approach than this, we would not be able to carry out the required calibrations within a time-frame which is realistic for the industry.

Overall, it is clear that the calibration philosophies described here are much more consistent with the theoretical foundations than the the day-by-day approach and more flexible than the rigid calibration. The question next is how these approaches can be extended to cater for multiple countries as well as exchange rates and then how good the fit is.

We have already remarked that because there is no common general testing framework, it is hard to make comparisons between the results of our fits and those obtained using other models. However, in §6.8 we will attempt to give a guide to the quality of our fits against some of the others quoted in the literature.

6.6 Modelling Exercise 4: fitting multiple countries

One of the many appealing features of the potential approach is the ease with which yield curves in many countries can be handled. Once we have defined our Markov chain and the yield curve for one country in terms of a function of that Markov chain, we can extend the model to include other countries without the need of any new sources of randomness. All that is required is a new function of the Markov process.

Therefore, for each country j in the model, we take a new positive function g^j and positive real α^j and define the state-price density for this country by

$$\zeta_t^j = R_{\alpha^j} g^j(X_t). \tag{6.30}$$

The underlying Markov chain is the same for all countries. Each g^j and α^j are included in the parameter vector θ which we use in the minimisation.

6.6.1 CI/RW calibrations for two and three country models

In this section we investigate two and three country fits using the CI and RW calibration procedures. We do not consider multi-country day-by-day or rigid calibrations because the work of §6.5 has clearly indicated that the CI and RW procedures are the most suitable calibration procedures.

For the two country fits we use USD and GBP data, again considering the period from 5th October 1994 to 6th March 1995. The three country fits will involve USD, GBP and DEM data from the same period. We use an 11-state Markov chain with a fully data imposed underlying structure.

- *Method:* CI and RW calibrations using 11-state Markov chain structures.
- *Markov structure:* data imposed structure.
- *Data used:* for two countries, USD and GBP, for three countries DEM is also included.
- *Parameters in minimisation, θ:*

  g vector and α for each country, upper triangular part of A, π.

6.6.2 Results

Tables 6.6 and 6.7 give summary statistics for the two and three country fits respectively. Further diagnostics for the two country fit can be found in Figures 6-14 and 6-15.

Although the results are not as good as the single country fits, they are still very promising. The median fitting error for the two country RW case with parameter $\beta = 0.2$ is around 3.5-4.5 bp per maturity. Notice also that the USD fit is consistently slightly worse than the fit for GBP. There seems to be little justification for this, but we are confident that with more effort and sensible choices of weights, this difference can be eliminated. For example, the boxplots of mod residuals for the USD CI calibration in Figure 6-14 indicate that the final maturity (corresponding to 10 year ZCB's) is considerably harder to fit than all the others. In order to pull this error down, we could try taking more states in the Markov chain, or try to alter the importance weighting of the maturity using the covariance matrix V. At present each maturity is weighted equally but it would perhaps be more sensible to adopt a weighting scheme which is dependent on the maturity. We could bias the error so that

$$\% \text{ error} = T_i(y_n(T_i) - Y(T_i, x; \theta))$$

for $i = 1, .., 8$, where T_i is the maturity.

Presently the fit for the short maturities is excellent, however, by adopting the above strategy we may find that the quality of the fit at the short end will deteriorate. Nevertheless, this method should help balance the fitting error over the yield curve.

CI/RW calibration statistics for two country fit (all values are in basis points)							
USD FIT							
β	Mean	Std. Dev.	Min	Q1	Median	Q3	Max
1.0 (CI)	93.082	36.671	28.606	60.205	87.108	127.584	168.642
0.8	59.183	19.086	16.899	45.524	55.416	71.192	123.935
0.6	47.947	13.160	18.402	38.282	47.458	54.798	91.380
0.4	43.583	11.195	16.978	36.625	42.664	50.295	74.498
0.2	38.153	11.564	15.619	30.589	37.258	45.501	73.653
0.1	36.775	11.586	15.335	29.052	35.939	44.924	73.549
GBP FIT							
β	Mean	Std. Dev.	Min	Q1	Median	Q3	Max
1.0 (CI)	62.968	26.815	9.552	41.592	62.022	81.679	128.058
0.8	37.816	13.280	9.292	30.283	37.566	46.950	76.318
0.6	33.551	12.923	8.444	24.170	33.806	41.459	76.651
0.4	31.264	12.261	12.330	22.219	28.115	38.620	74.929
0.2	29.184	12.738	6.061	19.819	27.085	38.287	73.427
0.1	28.338	12.902	7.086	19.283	26.448	37.447	72.702

Table 6.6: Summary statistics for the two country CI/RW fits, period 14 of USD and GBP. This table gives breakdowns for the GBP and USD fits individually.

Again we observe that the parameter movements over the fitting period are small, the posterior distribution (given in Figure 6-14) also has very different characteristics to that of the single country fit with more of the states being attached some initial weight. This trend is repeated in the posterior distribution for the three country RW fit (see Figure 6-16) where we see nearly all the states contribute to start with.

The RW plots in the three country case demonstrate that adding the third country does not substantially alter the quality of the fits, there is a deterioration but we still obtain fits of the order of 3.5-4.5 bp per maturity.

Diagnostic plots for two country (USD & GBP) CI calibration

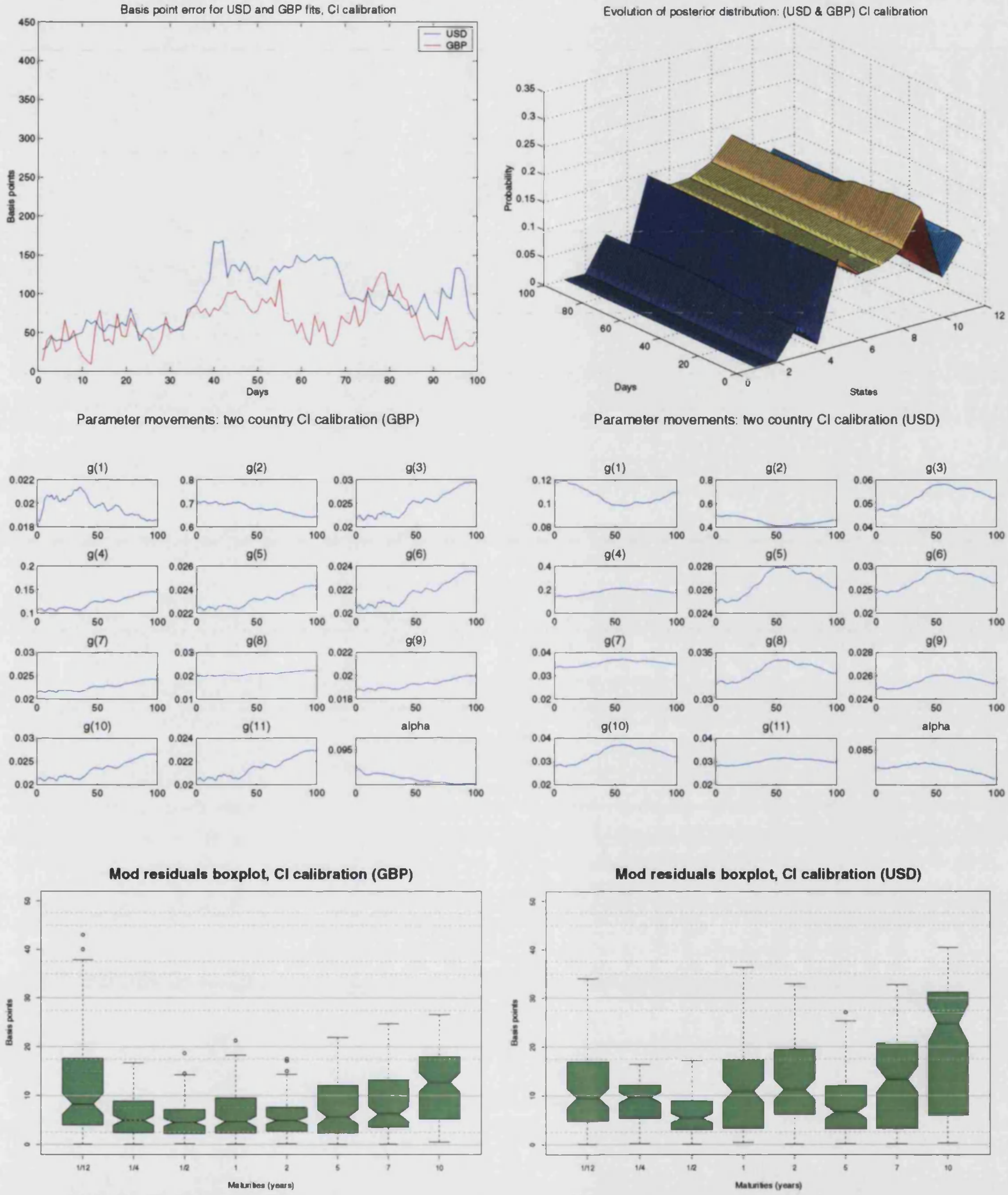

Figure 6-14: These plots refer to the two country fits for the CI calibration, USD and GBP. We show the basis point error plots for both countries (top left). The worst fit (blue) is the USD economy. The evolution of the **g** vector and α scalar are also given for each country. Again we normalise the **g** vectors using (6.12). The two boxplots are of the mod residuals for the USD and GBP fits.

Diagnostic plots for two country (USD & GBP) RW calibration ($\beta = 0.2$)

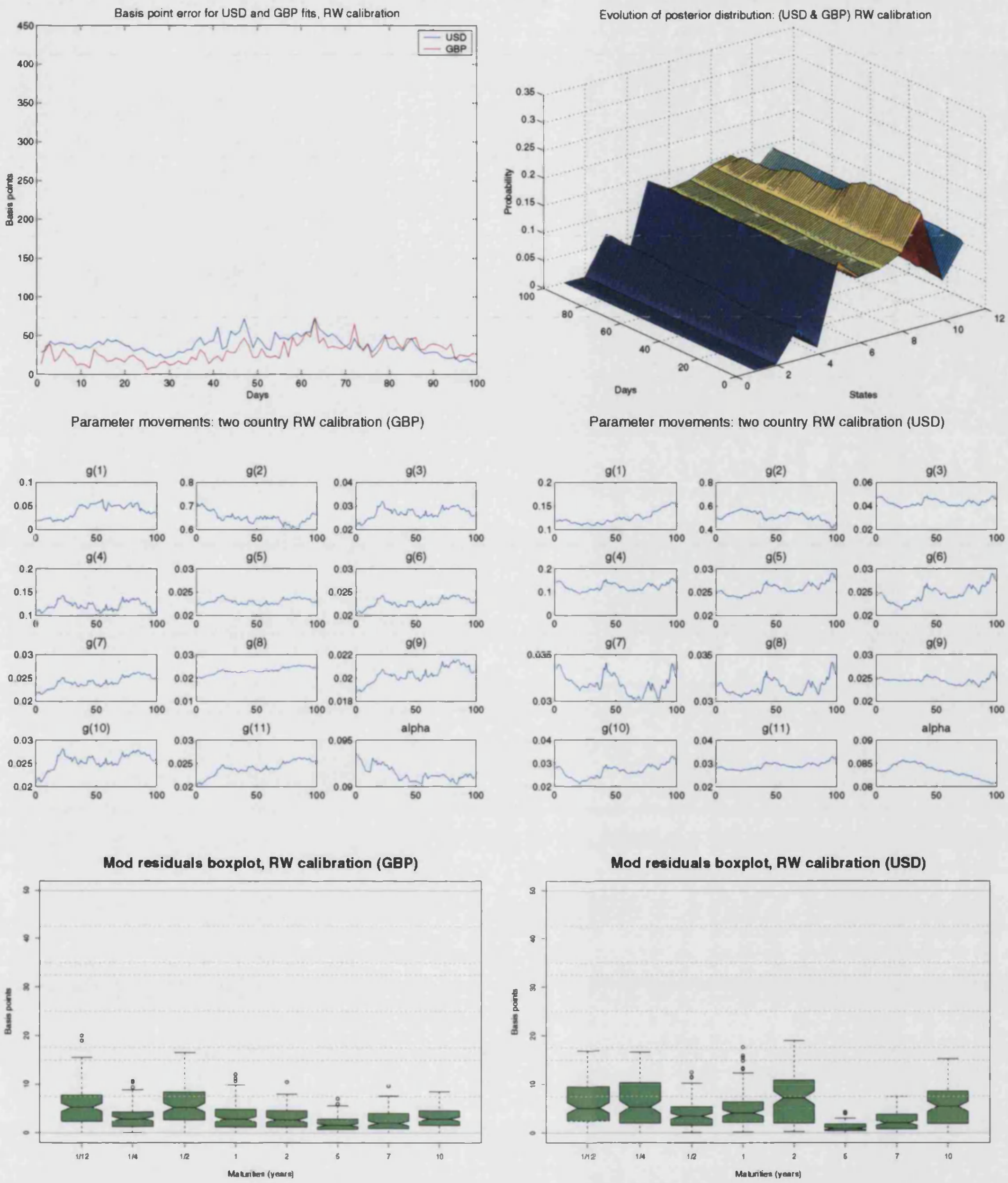

Figure 6-15: These plots refer to the two country fits for the RW calibration ($\beta = 0.2$), USD and GBP. We show the basis point error plots for both countries (top left). The worst fit (blue) is the USD economy.The evolution of the $\mathbf{g}$ vector and α scalar are also given for each country. Again we normalise the $\mathbf{g}$ vectors using (6.12). The two boxplots are of the mod residuals for the USD and GBP fits.

CI/RW calibration statistics for the three country fit
(all values are in basis points)

USD FIT							
β	Mean	Std. Dev.	Min	Q1	Median	Q3	Max
1.0 (CI)	115.678	32.023	48.308	100.543	115.006	136.413	211.331
0.8	86.260	22.713	34.222	66.194	86.097	106.046	130.637
0.6	67.065	20.279	14.589	52.759	63.658	83.099	113.933
0.4	50.983	14.494	23.901	41.129	50.661	57.949	89.969
0.2	39.511	10.414	21.605	30.209	39.640	47.340	67.369
0.1	35.153	10.444	17.487	27.883	33.912	42.666	66.016

GBP FIT							
β	Mean	Std. Dev.	Min	Q1	Median	Q3	Max
1.0 (CI)	84.958	29.853	28.229	60.984	81.629	107.327	149.090
0.8	45.772	13.036	12.617	36.644	46.403	52.012	82.245
0.6	39.031	13.808	12.529	31.140	37.188	46.512	87.466
0.4	35.006	12.430	13.078	27.435	33.425	40.875	76.725
0.2	31.140	10.532	12.523	24.456	29.408	37.529	70.204
0.1	29.123	10.458	7.467	22.943	27.858	35.956	67.502

DEM FIT							
β	Mean	Std. Dev.	Min	Q1	Median	Q3	Max
1.0 (CI)	66.798	21.886	32.260	50.378	62.351	79.501	132.468
0.8	49.857	10.955	30.902	41.828	47.704	56.204	86.001
0.6	45.070	9.925	28.391	38.174	43.194	50.179	74.346
0.4	40.557	8.014	25.880	35.359	40.172	45.085	72.847
0.2	37.050	6.360	23.124	32.613	37.530	41.188	59.138
0.1	35.640	5.704	23.380	32.013	34.881	39.556	55.702

Table 6.7: Summary statistics for the 100 day, three country CI/RW fits on period 14. This table gives breakdowns for USD, GBP and DEM.

Diagnostic plots for three country (USD, GBP & DEM) RW calibration ($\beta = 0.2$)

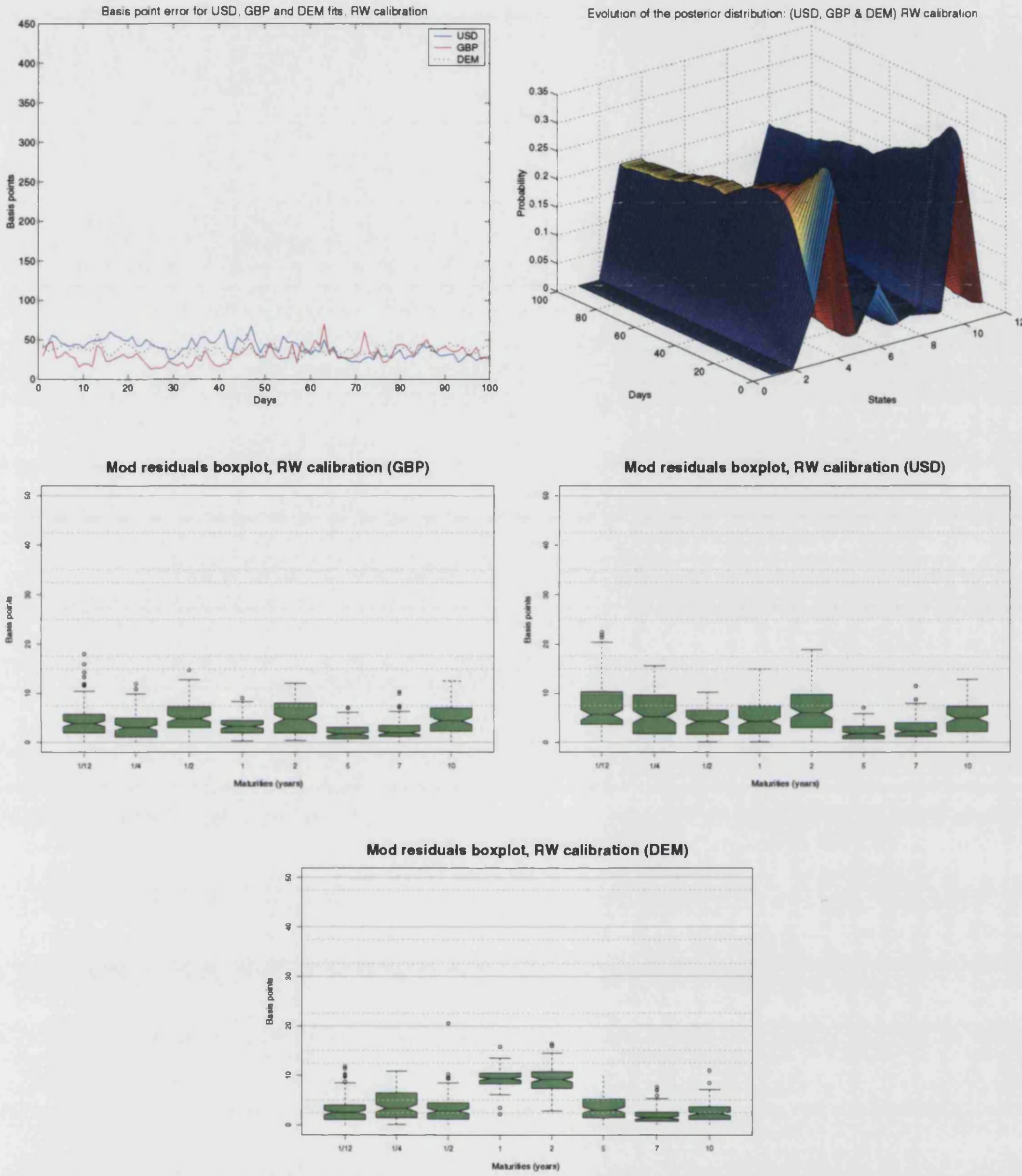

Figure 6-16: These plots are for a three country fit using the RW calibration method with $\beta = 0.2$. The first plot shows the basis point error for each of the three countries. The worst fit (blue line) is achieved by the USD and the the best fit (red line) is the GBP. The second plot (top right) shows the evolution in the posterior distribution during the fitting process. The boxplots are of the mod residuals for each maturity.

6.6.3 Discussion

Extending our conditional-independence and random walk calibrations of §6.5 so that they include the term structure of more than one country has proved to be a straightforward exercise. This is clearly an advantage that the Markov chain potential model has over some of the other models that have been proposed in the literature.

6.7 Modelling Exercise 5: foreign exchange rates

In this the final modelling exercise, we will investigate a method which will allow us to introduce exchange rates into our calibration frameworks. As we observed in §4.4, one of the most appealing features of the potential approach is that we can fit exchange rates without having to introduce any new sources of randomness. Therefore, our aim here will be to create a model which is driven by a single Markov chain structure and can handle the term structure between multiple countries along with the exchange rates between them.

Let us define

ζ_t^i as the state-price density for country i's assets at time t

and

$c^{ij}(t)$ as the observed value of 1 unit of country j's currency in terms of country i's currency at time t.

Rogers [84] shows, under certain assumptions (satisfied in the complete markets case),

$$\frac{c^{ij}(t)}{c^{ij}(0)} = \frac{\zeta_t^j}{\zeta_t^i} \quad \text{with} \quad \frac{\zeta_0^j}{\zeta_0^i} = 1. \tag{6.31}$$

If the model were correct then the value of the observation $c^{ij}(t_n)$ would be $C^{ij}(X_n;\theta)$, where X is the underlying Markov process and

$$C^{ij}(X_n;\theta) = \frac{e^{-\alpha_j t_n} R_{\alpha_j} g_j(X_n)}{e^{-\alpha_i t_n} R_{\alpha_i} g_i(X_n)} \tag{6.32}$$

To incorporate the exchange rate into our framework, we pick a base currency (the 'domestic currency') in relation to which exchange rates for all other currencies in the model will be defined. In the work which follows, we will let the base country be country 1.

6.7.1 Incorporating exchange rates into the rigid calibration

Although we do not intend to present any numerical results for the rigid calibration approach with multiple countries and exchange rates, let us still begin by demonstrating how exchange rates could be included within this calibration procedure if so desired.

We would again calibrate our model to a period of K days data, and minimise using

$$\sum_{j=1}^{K}[b(y_j - Y(1;\theta)) + q_1 s_j] - \log f_0(\theta), \tag{6.33}$$

where $-q_1$ is the diagonal entry in the first position of the Q-matrix Q.

Our vector θ will include g_j and α_j for all countries j together with the other usual parameters relating to the underlying structure. The difference when including exchange rates is that we normalise the g-vectors so that the model generated exchange rate for the final day in the calibration period is exactly correct. This is done by imposing an arbitrary normalisation for g_1 (the g-vector for country 1):

$$\sum_{k\in\mathcal{X}} g_1(k) = 1,$$

and for the g_j of the other countries (noting that the initial state is chosen to be state 1) we normalise so that

$$c^{1j}(d_K) \equiv \frac{R_{\alpha_j} g_j(1)}{R_{\alpha_1} g_1(1)},$$

where d_K is the last day in the calibration period. Having found our calibration vector θ^*, we test the model out-of-sample, fitting the yield curves for the days after the calibration period by only allowing changes in the posterior distribution of X. We compute the posterior distribution, π_n by using the recursive formula

$$\pi_n(x, \mathrm{y}_n) \;\propto\; \sum_{\xi} \pi_{n-1}(\xi, \mathrm{y}_{n-1}) p_{\xi x}(s_n; \hat{\theta}_n) \exp[-b(y_n - Y(x; \hat{\theta}_n))]$$
$$. \exp[- \sum_{\forall j \in G} b(c_n^{1j} - C^{1j}(x; \hat{\theta}_n))], \qquad (6.34)$$

where G is the set of countries excluding the base country. The proportional sign is because the above is subject to normalisation; we insist $\sum_{i=1}^{N} \pi_n(X_n(i), \mathrm{y}_n) = 1$ for all n. The initial distribution, π_0 is a point mass on the state chosen as the initial state in the calibration exercise. We calculate derivative prices in the same way as was described in §6.4 and the model's exchange rate t_n days after the calibration period is given by

$$\begin{aligned} C^{1j}(\hat{\theta}_n) &= \mathbb{E}\left[\frac{e^{-\alpha_j t_n} R_{\alpha_j} g_j(X_n)}{e^{-\alpha_1 t_n} R_{\alpha_1} g_1(X_n)} | \mathcal{F}_{t_n} \right] \\ &= \sum_{x \in \mathcal{X}} \pi_n(x, y_n) \frac{e^{-\alpha_j t_n} R_{\alpha_j} g_j(x)}{e^{-\alpha_1 t_n} R_{\alpha_1} g_1(x)}. \end{aligned}$$

6.7.2 Incorporating exchange rates into the CI and RW calibrations

Let us now explain how to incorporate the exchange rate into the CI and RW calibration procedures. In this case, the vector of parameters θ will now include g_j and α_j for all countries j together with the other parameters usually included. At each stage we will normalise the g-vectors. This is done by normalising g_1 (the g-vector for country 1) using

$$\sum_{k \in \mathcal{X}} g_1(k) = 1,$$

and the g_j of the other countries, we normalise using the initial posterior $\pi_0(\cdot)$, so that

$$\begin{aligned} c^{1j}(0) &\equiv \mathbb{E}\left[\frac{R_{\alpha_j} g_j(X_0)}{R_{\alpha_1} g_1(X_0)} | \mathcal{F}_{t_0} \right] \\ &= \sum_{x \in \mathcal{X}} \pi_0(x, y_0) \frac{R_{\alpha_j} g_j(x)}{R_{\alpha_1} g_1(x)}. \end{aligned}$$

The likelihood function (6.19) becomes

$$\begin{aligned} L_n(x, \mathrm{y}_n, \theta) \;\propto\; & \sum_{\xi} \pi_{n-1}(\xi, \mathrm{y}_{n-1}) p_{\xi,x}(s_n; \theta) \exp[-b(y_n - Y(x; \theta))] \\ & . \exp[-\sum_{j \in G} b(c_n^{1j} - C^{1j}(x; \theta))] \exp[-\frac{1}{2}(\theta - \hat{\theta}_{n-1}) S_{n-1}^{-1} (\theta - \hat{\theta}_{n-1})]. \end{aligned} \tag{6.35}$$

where G is the set of all countries excluding the base country. We again sum this expression over x, and numerically pick θ which maximises it; the maximising value is our new estimate $\hat{\theta}_n$ of θ. We compute the posterior distribution, π_n by using the recursive formula

$$\begin{aligned} \pi_n(x, \mathrm{y}_n) \;\propto\; & \sum_{\xi} \pi_{n-1}(\xi, \mathrm{y}_{n-1}) p_{\xi x}(s_n; \hat{\theta}_n) \exp[-b(y_n - Y(x; \hat{\theta}_n))] \\ & . \exp[-\sum_{j \in G} b(c_n^{1j} - C^{1j}(x; \hat{\theta}_n))]. \end{aligned} \tag{6.36}$$

The model's exchange rate at time t_n will be given by

$$\begin{aligned} C^{1j}(\hat{\theta}_n) &= \mathbb{E}\left[\frac{e^{-\alpha_j t_n} R_{\alpha_j} g_j(X_n)}{e^{-\alpha_1 t_n} R_{\alpha_1} g_1(X_n)} | \mathcal{F}_{t_n} \right] \\ &= \sum_{x \in \mathcal{X}} \pi_n(x, y_n) \frac{e^{-\alpha_j t_n} R_{\alpha_j} g_j(x)}{e^{-\alpha_1 t_n} R_{\alpha_1} g_1(x)}. \end{aligned}$$

6.7.3 CI/RW calibration of a two country model with exchange rates

We are now in a position to investigate the fit using the CI and RW calibration approaches of term structures for the USD and GBP together with the exchange rate between the two. One motivation for investigating a model which can provide good fits for such a set up is that it would potentially allow us to price or hedge cross currency derivatives.

- *Method:* CI and RW calibration of two country and exchange rate using an 11-state Markov chain.
- *Markov structure:* data imposed structure.

- *Data used:* two countries, USD and GBP, together with dollar/pound exchange rate.
- *Parameters in minimisation,* θ*:*

g vector and α for both countries, π, upper triangular part of A.

6.7.4 Results

When we compare the basis point error plot for the fit in Figure 6-17 with the corresponding plot for the two country fit, Figure 6-14, we see similar characteristics.

Table 6.8 gives a breakdown of the statistics of the fit for different β parameters, we again focus on RW $\beta = 0.2$ and plots for this calibration procedure can be found in Figures 6-19 and 6-20. The results show that the median error is about 4.5-5.5 bp out per maturity and the exchange rate error is not more than five hundreths of a cent (0.5 bp).

CI/RW calibration statistics for two country and exchange rate fit
(all values are in basis points)

USD FIT							
β	Mean	Std. Dev.	Min	Q1	Median	Q3	Max
1.0 (CI)	146.245	54.529	38.552	101.450	158.334	190.084	310.810
0.8	82.139	29.645	35.550	60.063	78.209	105.595	144.969
0.6	72.696	29.291	30.794	51.712	62.452	98.970	138.709
0.4	58.471	20.738	21.819	43.719	52.419	71.981	115.074
0.2	49.979	16.727	19.442	37.926	47.491	60.373	96.033
0.1	42.505	12.762	5.670	34.353	42.572	48.860	82.313

GBP FIT							
β	Mean	Std. Dev.	Min	Q1	Median	Q3	Max
1.0 (CI)	103.685	41.776	16.543	72.962	101.316	129.850	211.00
0.8	48.495	17.366	10.098	39.003	47.939	58.941	86.957
0.6	42.928	14.936	13.318	32.262	40.831	53.962	83.303
0.4	32.557	12.410	10.285	22.825	32.933	40.758	74.011
0.2	35.910	12.144	13.076	27.088	34.235	43.552	79.165
0.1	30.402	11.581	10.049	21.721	29.326	36.612	73.550

Table 6.8: Summary statistics for the two country and exchange rate CI/RW fits over 100 days (period 14). This table has breakdowns for the USD and GBP fitting errors.

Diagnostic plots for two country (USD & GBP) CI calibration with exchange rates

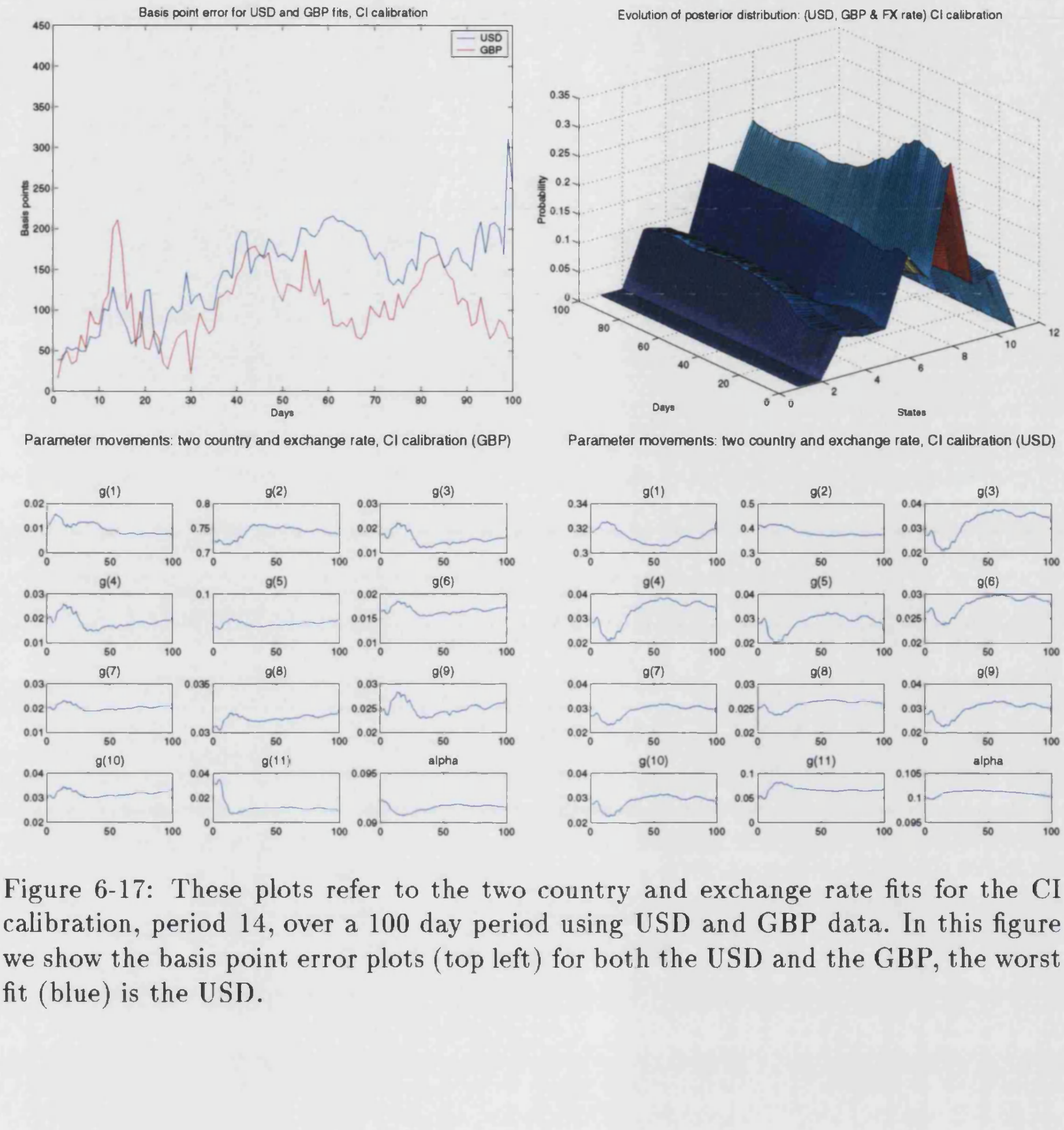

Figure 6-17: These plots refer to the two country and exchange rate fits for the CI calibration, period 14, over a 100 day period using USD and GBP data. In this figure we show the basis point error plots (top left) for both the USD and the GBP, the worst fit (blue) is the USD.

Diagnostic plots for two country (USD & GBP) CI calibration with exchange rates

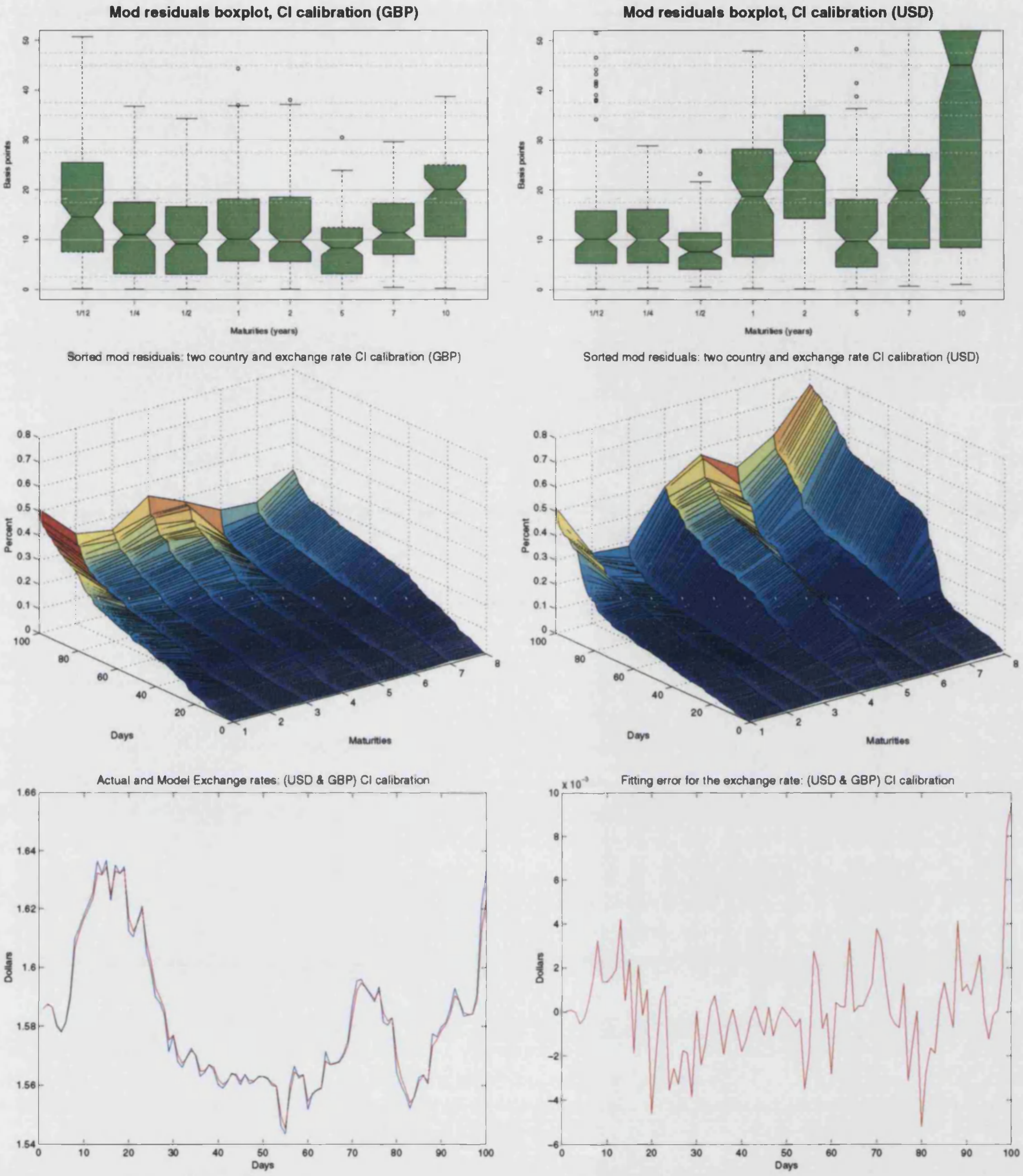

Figure 6-18: These plots refer to the two country and exchange rate fits for the CI calibration of the USD and GBP, period 14 data. The penultimate plot in this figure shows the observed data and the fitted curve for the exchange rates (there are two curves in this picture). The final plot is of the fitting error in the exchange rate.

Diagnostic plots for two country (USD & GBP)
RW calibration with exchange rates ($\beta = 0.2$)

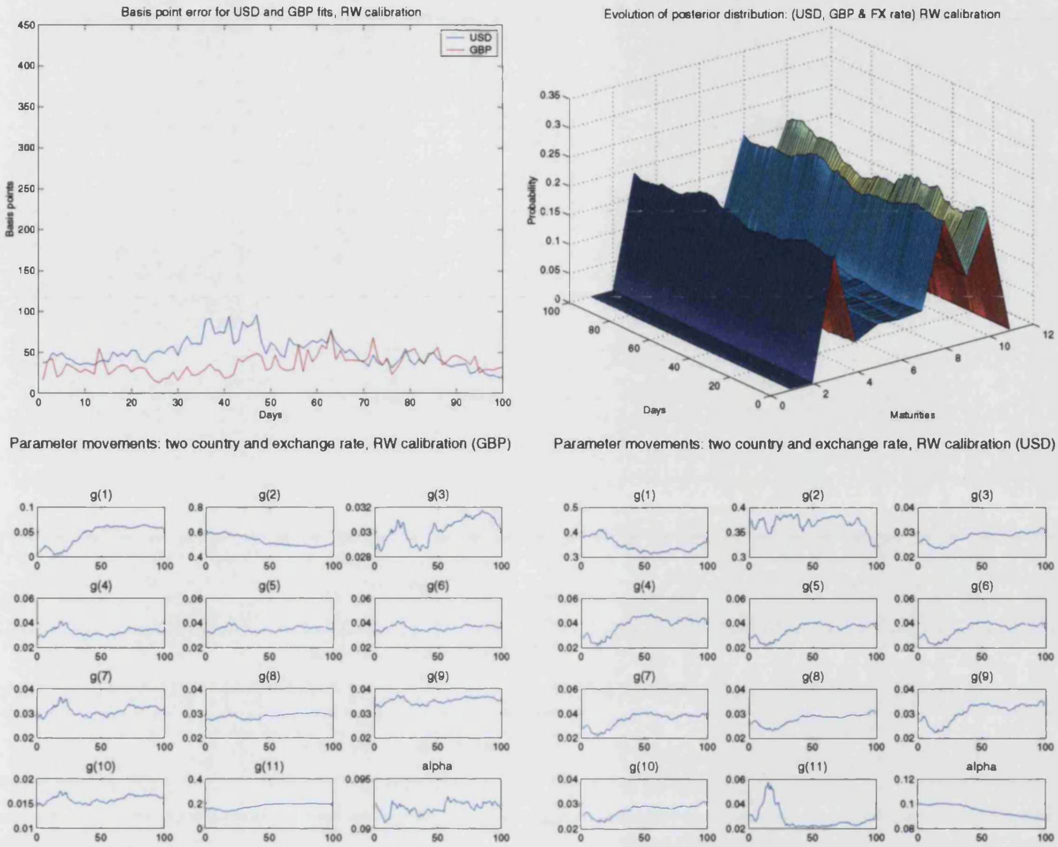

Figure 6-19: These plots refer to the two country and exchange rate fits for the CI calibration, period 14, over a 100 day period using USD and GBP data. In this figure we show the basis point error plots (top left) for both the USD and the GBP, the worst fit (blue) is the USD.

Diagnostic plots for two country (USD & GBP)
RW calibration with exchange rates ($\beta = 0.2$)

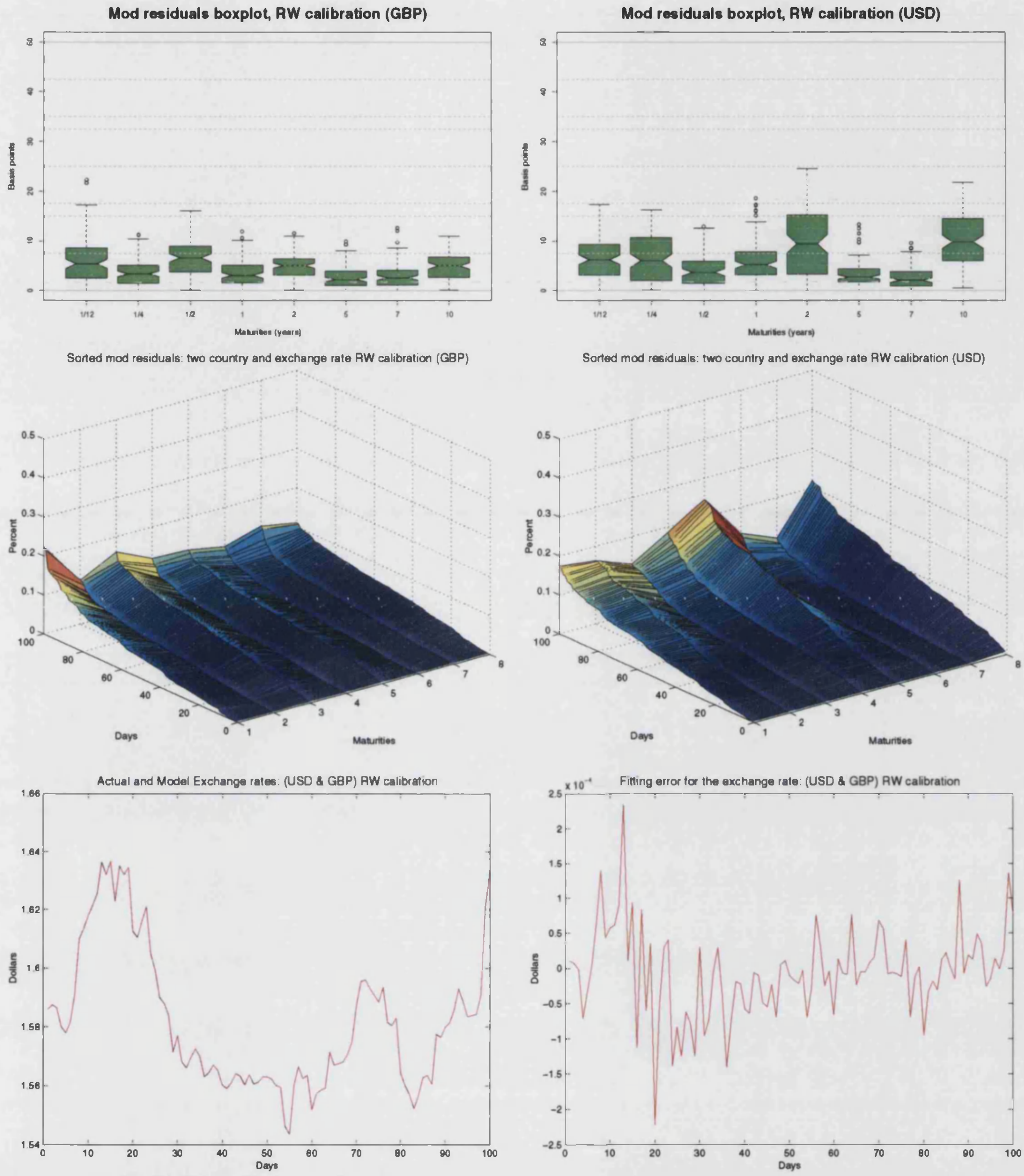

Figure 6-20: These plots refer to the two country and exchange rate fits for the RW calibration of the USD and GBP, period 14 data. The penultimate plot in this figure shows the observed data and the fitted curve for the exchange rates (there are two curves in this picture). The final plot is of the fitting error in the exchange rate.

6.7.5 Discussion

Incorporating exchange rates into our calibration procedures can be accomplished with little difficulty. Clearly the fits of the yield curve do suffer as a result and there is a payoff between the quality of the fit of the exchange rate and that of the yield curve. We also found that by attaching more weight to the yield curve and lowering it for the exchange rate we were able to obtain much better yield curve agreement with only a small deterioration in the exchange rate.

Although the quality of the fits obtained for multiple countries together with exchange rates are probably not good enough to trade off, it should be remembered that this is only really a pilot study. If we take into account the fact that we only have 11 states in our Markov chain, the omens are good for the quality of fit that might be achieved with much bigger chains.

6.8 Empirical results from other studies

Although we have already briefly discussed other empirical studies in §3.6, we have not mentioned the quality of fits achieved in those papers. Unfortunately, even a broad comparison of results appears difficult as studies such as those by Moraleda & Pelsser [78], Amin & Morton [14] and Bühler et al [30] adopt very different calibration procedures and consider different classes of interest rate models. The interested reader is therefore referred to the original papers to see the analysis and results for their investigations.

One paper we can directly compare our results with is that of Rogers & Zane [85] which we discussed in §4.5. This paper investigates calibrating a diffusion based potential model to the USD and GBP. They present a couple of different calibration techniques, one of which is similar to the day-by-day calibration which we implemented in §6.3. A more realistic approach they consider is one they term a quadratic constrained fit. Figure 6-21 shows their results for this fitting approach[5].

[5]The author wishes to thank Prof L. C. G. Rogers for allowing us to reprint these plots

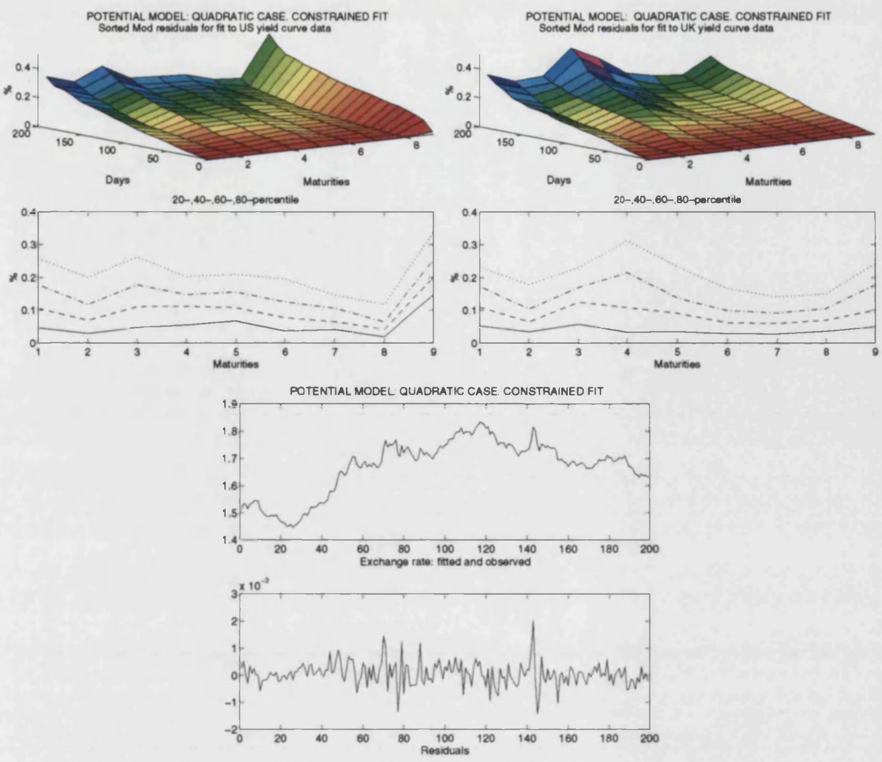

Figure 6-21: Three plots from Rogers & Zane [85] showing their two country and exchange rate fits.

Their fits have been conducted over a fitting period of 200 days, nevertheless if we compare them with the two country and exchange rate results obtained using our RW calibration ($\beta = 0.2$) in Figures 6-19 and 6-20, we see that our results compare very favourably.

6.9 Conclusions, Criticisms and Discussion

Our study of Markov chain potential models has yielded a number of interesting results. Although this modelling regime has not yet been considered by the industry and is very different to the ones currently used, it appears to have several appealing features;

- it provides a quick and straightforward means of pricing complicated derivatives. The traditionally difficult integration over multi-dimensional spaces which is often needed when using diffusions as the underlying process reduces to an easy-to-handle sum over states in the Markov chain framework;
- it can model the yield curves from any number of countries without the need for

added sources of randomness and without complicating the underlying Markov process;

- we can make rational choices for the underlying Markov structure based on market data;
- modelling exchange rates within the same framework is a straightforward extension;
- it also follows that this model is well suited to pricing cross-currency derivatives.

The work in this thesis serves as a pilot study into the feasibility of this approach. We begin in Chapter 5 by presenting a theoretical result which demonstrates how by choosing the underlying structure in the potential approach to be an Ornstein Uhlenbeck process, it is possible to replicate the bond prices given by Vasicek's spot rate model. The work in this chapter has been solely concerned with issues of calibration of Markov chain potential models to market data. The aim has not been to compare this model with others but more to present and implement different ways of calibrating the model to the term structure and exchange rates of several countries.

Most of the trials carried out have been restricted to Markov chain models with only 11-states and this would undoubtedly be much bigger in practice. The aim throughout the study has been to investigate computationally feasible models, by this it is meant that the models proposed should be able to be used in a trading environment. The longest computer runs we conduct here are those of the conditional-independence fits over 100 days. Each run of the single country fit for this period took of the order of 200 cpu minutes and the two country with exchange rate fit was of the order of 300 cpu minutes[6]. Clearly these figures are dependent on the starting point for the minimisation routines and the convergence criteria imposed.

The reasons why we choose not to quote exact computation times for our runs were twofold. Firstly, given the current rate of advancement of computer technology it is likely that any figures we might give would be outdated in a very short period of time. Indeed, the computer technology currently available to us (in this academic environment) are multiuser machines which are not close to the cutting edge technology now on the market. A second reason for not giving times is that much of the numerics involved in this study would be well suited to parallelization and if the industry were to adopt these models this is the route they would almost certainly choose to go down.

[6]These computation times were for runs conducted on a Sun Ultra E3500 with 400 MHz UltraSPARC II processor

The results themselves have been promising given that we are insisting on fitting a time-homogeneous model with such a small statespace. We found that when allowing a different model to be fitted each day (day-by-day calibration), we were able to fit a single yield curve with median errors of the order of 1 bp per maturity. Conversely, the rigid calibration of fitting 5 days of data and then using the calibrated model to fit subsequent days yielded median errors of the order of 6 bp per maturity.

The best method appeared to be one which was between these two extremes, that is the conditional-independence and random walk calibrations. We were able to produce single-country fits with median errors of around 2.5 bp per maturity with good parameter stability. Obviously, incorporating more than one country inevitably worsened the fit; when we fitted USD and GBP data, the median error was around 3.5-4.5 bp per maturity, adding a third country (DEM) increased the error slightly. The inclusion of the exchange rate in the two country fit gave errors of the order of 4.5-5.5 bp which is high for trading. However, it is again worth emphasizing the severe restrictions imposed and the fact that this is the error for modelling two economies together with the exchange rate between them.

There are obvious extensions which could be carried out, some of which are mentioned in the avenues for further research section which follows this one. We are confident that the fit could be improved with more states in the chain and by focusing more on the weights of each maturity. One other avenue to pursue would be incorporating derivatives other than ZCBs in the calibration/fitting process and investigating how this affects the fit.

As regards the question of whether the industry would use such a model, we have already mentioned many of the appealing features. However, one of the main stumbling blocks is that this model is clearly very different to the ones currently used by the industry and practitioners may well be reluctant to drastically change their existing models unless the perceived benefits of a new approach were very high.

6.10 Avenues for further research

The work in this thesis has opened up a number of interesting avenues for further research. Along with the possible extensions to the theory which we have suggested in the body of the text, it would be interesting to extend the framework so that it is capable of hedging derivatives.

It would also be of interest to investigate how well the existing framework could price more complicated cross-currency derivatives. One pricing problem which could be looked at is that of pricing a cross-currency swaption. Although we will not investigate this here, it is fairly straightforward to obtain expressions for prices of such derivatives in this modelling framework. All that is required is good quality market data.

Part III

Exogenous Shocks to a Market Economy

Chapter 7

Examining the Effects of Shocks to a Market Equilibrium

7.1 Introduction

Events which influence the financial world are not always expected and the effect they can have is often far from predictable. In this and the next chapter we present a classical equilibrium model and study the changes which happen when we impose a sudden unexpected shock to the market equilibrium.

Sudden changes and moves in the world markets are not uncommon, for example in recent years we have had; the stock market crash of October 1987, the sterling crisis of 1992, the collapse of the Korean economy at the end of 1997 and the Russian default in August 1998. These events result in large market movements and are often characterised by a lack of liquidity in the critical markets. The effects can potentially be far more important than many of the risks usually described as market risk.

The structure of the model we use in this study is quite conventional. Under the assumption of complete markets, we construct an economy which has multiple investors with heterogeneous preferences who may invest in a small number of productive risky assets and a riskless deposit account. Each investor aims to maximise their expected integrated utility of consumption over all future time (this is a continuous time, infinite horizon model). Each risky asset has a dividend stream which we model as a log Brownian process and agents have so-called power-law or constant relative risk aversion utility function.

We examine the effect a sudden shock has on the economic equilibrium. The shock is imposed via the dividend stream of the risky assets and a comparison of the equilibrium before and after the event is made. We make no attempt to describe the trajectory of the economy through the period of change. This analysis may offer some insight into certain "real-world" situations. For example it may give us an indication of what would happen if a large but volatile market participant such as the Russian or Asian economy suffered a severe downturn in fortunes.

There are now many papers in the economic and finance literature concerned with equilibrium modelling. Much of the initial work on optimal consumption-investment policy was carried out by Merton [73] [74] [75] who studied the single agent optimal control problem. Merton produced closed form expressions for the consumption and investment policies when the utility function of consumption, $U(c)$, was of the HARA (Hyperbolic Absolute Risk Aversion) class and satisfied $U'(0) = \infty$.

Among the other early well known models is the theory developed by Lucas [71] for equilibrium asset pricing in a discrete-time setting, this model provides the flexibility of allowing us to pick an exogeneous consumption-portfolio plan.

The starting point for a large number of continuous-time equilibrium models for the valuation of contingent claims and the term structure of interest rates is the Cox, Ingersoll and Ross (CIR) [36] valuation formula. Many of the studies which use this model assume log utility functions for tractability.

The work of Cox & Huang [35], Karatzas, Lehoczky, Sethi & Shreve [64] and Karatzas, Lehoczky & Shreve [65] [66] were an important breakthrough in the study of this kind of optimal consumption-investment type problem.

The first of these papers was Karatzas, Lehoczky, Sethi & Shreve [64]. In their work the authors obtained explicit formulas for optimal single agent consumption and investment policies when the stock price process is geometric Brownian motion. They were also able to remove the restriction that $U'(0) = \infty$.

A martingale based characterization for the optimal decisions of a single agent under a general class of stock price processes with more general utility functions was then given in Karatzas, Lehoczky & Shreve [65]. The same methodology was independently developed and applied to diffusion models by Cox & Huang [35]. Finally, the single agent equilibrium was extended to a multi-agent problem in Karatzas, Lehoczky & Shreve [66]; results on existence and uniqueness of equilibria solutions were given in this paper.

There are now a number of related studies which use and extend the work of the above authors. However, much of the work is concerned with proofs of existence and the formulation of closed form expressions. There are still very few studies which examine the properties of particular equilibrium solutions and give a description of the behaviour or characteristics exhibited. The work of Dumas [43] and that of Wang [91] are notable exceptions to this. Their interesting studies begin to consider the equilibrium properties for particular two-investor models with specially chosen heterogeneous investors.

In this part of the thesis we shall consider an equilibrium model which has a similar construction to that discussed for the two-investor case in Wang [91]. This is a conventional intertemporal equilibrium consumption-investment model and it has elements in common with the classical Merton [74] model and the work in the seminal paper of Breeden [23]. The model in Dumas [43] also has close links with our model. We will examine the behaviour and characteristics of an equilibrium which has multiple investors who are heterogeneous in risk and impatience preferences. The paper of Wang [91] together with that of Karatzas et al [66] may be consulted for questions concerning the existence and uniqueness of equilibria for this model.

Wang considers two classes of investors and in analysing the equilibria he imposes the simplifying restriction that investors in class 1 have logarithmic utility and investors in class 2 have a square root utility function. We will consider a more general situation and numerically investigate equilibria with four heterogeneous agents each with CRRA utility. We examine the change in portfolios of the investors, the affect on stock prices and the variation in interest rates as a result of shocks to the equilibrium solution.

One motivation for attempting to examine the equilibrium characteristics is that it allows us to understand what might happen after a shock to the market. The literature has few papers which genuinely analyse the effect of external shocks to an equilibrium solution as we will attempt.

There are, however, several studies which focus on policy measures and the role of liquidity in precipitating financial crashes. For example Diamond and Dybvig [40] discuss how illiquidity and bank runs can cause real economic damage; they look at contracts and government provision which may prevent this. Gennotte & Leland [46] consider how hedging strategies can trigger liquidity concerns which, in turn, can lead to sharp market falls. Short-term borrowing has also often been blamed for such crises, the reader is referred to Diamond & Rajan [41] for an account of this. Masson [72] develops a simple two-country balance of payments model to illustrate the role of contagion (also see Allen & Gale [13]) in the collapse of markets; discussing in particular the recent turmoil in Asia. Further economic and political analysis of the Asian crisis may be found in Mishkin [77]; this paper suggests causes as well as strategies to prevent collapses.

Finally, the paper of Blejer, Feldman and Feltenstein [19] develops a framework to analyse how an exogenous shock affects an initially solvent banking system. Their focus is to use a simple general equilibrium model and analyse the potential economic consequences and policy responses when sudden shocks affect bank deposits in the economy.

The plan for this and the next chapter is as follows. We start in §7.2 by setting-up the problem, explaining the assumptions we make and the terminology used. In §7.3 we derive the system of equations which must be satisfied by the equilibrium solution giving expressions for the asset price and the wealth of each investor. We also give a sufficient condition for convergence of our system of equations. The interest rate for the riskless deposit account is determined endogenously and in §7.4 we derive an explicit expression for this rate. We end this chapter with a few remarks, §7.5.

Chapter 8 follows on directly from the work of this chapter. We present a numerical procedure to synthesize shocks to the market and investigate two explicit examples providing the results and analysis of our study.

7.2 The Set Up

Consider a world with J investors, M assets and a riskless deposit account with endogeneously determined interest rate $(r_t)_{t\geq 0}$.

Each asset i has a dividend stream $(\delta_i(t))_{t\geq 0}$ which we model as a log Brownian process in the form

$$dX_i(t) \equiv d(\log \delta_i(t)) = \sigma_i dW_t^i + \mu_i dt \tag{7.1}$$

where $(W_t^i)_{t\geq 0}$ is a standard Brownian motion (independent for all i), μ_i is the mean rate of return and σ_i is the volatility of the asset. We define $\delta(t) = (\delta_1(t), .., \delta_M(t))$ and also $\Delta(t) = 1.\delta(t) = \sum_{i=1}^{M} \delta_i(t)$.

Each investor j has a utility function, $U_j(t,c)$, given by

$$U_j(t,c) = e^{-\rho_j t}\frac{c^{1-R_j}}{(1-R_j)}, \tag{7.2}$$

where c represents the consumption rate of the agent, ρ_j is the rate of impatience of agent j and $R_j > 0$ ($R_j \neq 1$) for all j.[1] Investors seek to

$$\max_{c_j}\left\{\mathbb{E}_0\left[\int_0^\infty U_j(t,c_j(t))dt\right]\right\} \tag{7.3}$$

subject to the constraint

$$w_j(t) \geq 0 \tag{7.4}$$

for all t. Here $w_j(t)$ is the wealth and $c_j(t)$ is the consumption rate of agent j at time t. $\mathbb{E}_t \equiv \mathbb{E}[\cdot|\mathcal{F}_t]$ is the $\mathbb{P}$ measure conditional expectation given the information $\mathcal{F}_t$ available at time t.

Agent j invests an amount $\phi_j(t)$ in the riskless bank deposit account and holds $\pi_j^i(t)$ in share i ($i = 1,..,M$) at time t. We denote agent j's portfolio of holdings in the risky assets by the vector

$$\pi_j(t) \equiv (\pi_j^1(t),...,\pi_j^M(t))^T.$$

Therefore, it follows that the time t wealth of agent j is

$$w_j(t) = \pi_j(t).S_t + \phi_j(t) \tag{7.5}$$

where S_t is the vector of prices at time t for the M assets.

Although we allow the risky assets to be sold short and money to be borrowed from the bank, we shall impose the restriction that there is exactly one share of each productive asset in existence, and cash is in zero net supply:

$$\sum_{j=1}^{J}\pi_j^i(t) = 1\ \forall t\ \ i=1,..,M \quad \text{and} \quad \sum_{j=1}^{J}\phi_j(t) = 0,\ \forall t. \tag{7.6}$$

[1]The case $R_j = 1$ is that of the logarithmic utility and can be considered analogously but we do not explicitly discuss it here.

We will assume throughout that we are in the complete market situation. Therefore, there exists a unique probability measure $\mathbb{P}^*$ (equivalent to $\mathbb{P}$) under which all discounted asset prices are martingales. The state price density, ζ is defined by

$$\zeta_t = \exp(-\int_0^t r_s ds) \frac{d\mathbb{P}^*}{d\mathbb{P}}\bigg|_{\mathcal{F}_t}.$$

Under these assumptions, it is well known (see, for example, Karatzas, Lehoczky & Shreve [66]) that for some constants[2] λ_j the marginal utilities of agents' consumption and the state-price density process $(\zeta_t)_{t\geq 0}$ are linked by

$$U_j'(t, c_j(t)) = \lambda_j \zeta_t, \tag{7.7}$$

where prime represents differentiation with respect to consumption. The share price is the net present value of all future dividends

$$S_i(t) = \zeta_t^{-1}\, \mathbb{E}_t \left[\int_t^\infty \zeta_u \delta_i(u) du \right] \tag{7.8}$$

and the wealth process is the net present value of all future consumption

$$w_j(t) = \zeta_t^{-1}\, \mathbb{E}_t \left[\int_t^\infty \zeta_u c_j(u) du \right]. \tag{7.9}$$

Finally to complete the first part of the story we also impose the restriction

$$\Delta(t) \equiv \sum_{i=1}^{M} \delta_i(t) = \sum_{j=1}^{J} c_j(t) \tag{7.10}$$

on the dividend payments of the assets and the consumption of the agents. This equation, (7.10), will be referred to as the market clearing condition.

7.3 Calculating the Equilibrium

Since we consider the situation where the utility of each agent j is of the form

$$U_j(t, c_j(t)) = e^{-\rho_j t} \frac{c_j^{1-R_j}(t)}{1 - R_j}, \tag{7.11}$$

[2]The constants λ_j are often referred to as Lagrange multipliers.

it follows that

$$U_j'(t, c_j(t)) = e^{-\rho_j t} c_j^{-R_j}(t). \tag{7.12}$$

Recall from §7.2 that under complete market assumptions we have $U_j'(t, c_j(t)) = \zeta_t \lambda_j$. Therefore,

$$c_j(t) = (\zeta_t \lambda_j e^{\rho_j t})^{-\frac{1}{R_j}}, \tag{7.13}$$

and so the market clearing (see (7.10)) gives

$$\Delta_t \equiv \sum_{i=1}^{M} \delta_i(t) = \sum_{j=1}^{J} c_j(t) = \sum_{j=1}^{J} (\zeta_t \lambda_j e^{\rho_j t})^{-\frac{1}{R_j}}. \tag{7.14}$$

Ideally we would like to be able to make ζ_t the subject of this equation but unfortunately, this is the best we can do in this set up. Indeed, the only way to determine ζ_t here is to use a numerical root finding routine and we discuss this more fully in §8.2.

7.3.1 Share price and wealth expressions

We have already given expressions for the share price and the wealth ((7.8) and (7.9) respectively). Although we would like to make these more explicit, it is not possible to progress any further at the general level, the best we can do is write the wealth of any agent j as:

$$\begin{aligned} w_j(t) &= \frac{1}{\zeta_t} \mathbb{E}_t \Big[\int_t^\infty \zeta_s c_j(s) ds \Big] \\ &= \frac{1}{\zeta_t} \mathbb{E}_t \Big[\int_t^\infty \zeta_s (\zeta_s \lambda_j e^{\rho_j s})^{-\frac{1}{R_j}} ds \Big] \\ &= \frac{1}{\zeta_t} \int_t^\infty \mathbb{E}_t [\zeta_s^{(1-\frac{1}{R_j})}] (\lambda_j e^{\rho_j s})^{-\frac{1}{R_j}} ds, \end{aligned} \tag{7.15}$$

and write the share price process for asset i as:

$$\begin{aligned} S_i(t) &= \frac{1}{\zeta_t} \mathbb{E}_t \Big[\int_t^\infty \zeta_s \delta_i(s) ds \Big] \\ &= \frac{1}{\zeta_t} \int_t^\infty \mathbb{E}_t [\zeta_s \delta_i(s)] ds. \end{aligned} \tag{7.16}$$

7.3.2 Expression for the payoff

In this case, an expression for the payoff, $V_j(t)$, at time t for any agent j can easily be derived. Using (7.11) and (7.13) we see that

$$\begin{aligned}
V_j(t) &= \mathbb{E}_t[\int_t^\infty U_j(s, c_j(s))ds] \\
&= \mathbb{E}_t[\left(\frac{1}{1-R_j}\right)\int_t^\infty e^{-\rho_j s}(\zeta_s \lambda_j e^{\rho_j s})^{(1-\frac{1}{R_j})}ds] \\
&= \left(\frac{1}{1-R_j}\right)\int_t^\infty e^{-\rho_j s}(\lambda_j e^{\rho_j s})^{(1-\frac{1}{R_j})}\mathbb{E}_t[\zeta_s^{(1-\frac{1}{R_j})}]ds \\
&= \left(\frac{1}{1-R_j}\right)\int_t^\infty \lambda_j(\lambda_j e^{\rho_j s})^{(-\frac{1}{R_j})}\mathbb{E}_t[\zeta_s^{(1-\frac{1}{R_j})}]ds \\
&= \left(\frac{\lambda_j}{1-R_j}\right)\zeta_t w_j(t).
\end{aligned}$$

7.3.3 Expression satisfied by parameters at equilibrium

Next we wish to find the system of equations that the model parameters must satisfy to be in equilibrium. To do this we will use the wealth equation given in (7.5)

$$w_j(t) = \sum_{i=1}^{M} \pi_j^i(t) S_i(t) + \phi_j(t). \tag{7.17}$$

So putting (7.15) and (7.16) into (7.17) gives

$$\int_t^\infty \mathbb{E}_t[\zeta_s^{(1-\frac{1}{R_j})}](\lambda_j e^{\rho_j s})^{-\frac{1}{R_j}}ds = \sum_{i=1}^{M}\pi_j^i(t)\int_t^\infty \mathbb{E}_t[\zeta_s \delta_i(s)]ds + \phi_j(t)\zeta_t$$

for $j = 1, .., J$. Therefore we have the J equations

$$\int_t^\infty \left\{\sum_{i=1}^{M}\pi_j^i(t)\mathbb{E}_t[\zeta_s\delta_i(s)] - \mathbb{E}_t[\zeta_s^{(1-\frac{1}{R_j})}](\lambda_j e^{\rho_j s})^{-\frac{1}{R_j}}\right\}ds + \phi_j(t)\zeta_t = 0, \tag{7.18}$$

which we solve for λ_j, $j = 1, .., J$.

This system of equations cannot be solved analytically, so the only way we can proceed is via numerical means. However, this is also far from trivial as we have a complicated infinite multi-dimensional integral to evaluate. Moreover, there is no guarantee that the integral (7.18) actually converges.

7.3.4 Convergence of (7.18)

It is important to try to obtain a sufficiency condition under which convergence of the system of equations given in (7.18) occurs.

First notice that

$$\begin{aligned}\mathbb{E}_t[\zeta_s^{(1-\frac{1}{R_j})}](\lambda_j e^{\rho_j s})^{-\frac{1}{R_j}} &\leq \sum_{j=1}^{J}\mathbb{E}_t[\zeta_s(\zeta_s\lambda_j e^{\rho_j s})^{-\frac{1}{R_j}}] \\ &= \mathbb{E}_t[\zeta_s\Delta_s] \qquad (7.19)\end{aligned}$$

where the equality above comes from the market clearing condition. Similarly, by defining $\overline{\pi}_j = \max_i \pi_j^i$, we can also obtain the inequality

$$\begin{aligned}\sum_{i=1}^{M}\pi_j^i(t)\mathbb{E}_t[\zeta_s\delta_i(s)] &\leq \overline{\pi}_j(t)\mathbb{E}_t[\zeta_s\sum_{i=1}^{M}\delta_i(s)] \\ &= \overline{\pi}_j(t)\mathbb{E}_t[\zeta_s\Delta(s)]. \qquad (7.20)\end{aligned}$$

Hence in view of (7.19) and (7.20), it follows that to determine conditions for convergence of (7.18) we need to ensure that $\int_0^\infty \mathbb{E}_t[\zeta_s\Delta_s]ds$ converges. To find conditions for this, first observe that

$$\Delta_t = \sum_{j=1}^{J}(\zeta_t\lambda_j e^{\rho_j t})^{-\frac{1}{R_j}} \leq \begin{cases} \zeta_t^{-1/\overline{R}}(\sum_{j=1}^{J}(\lambda_j e^{\rho_j t})^{-\frac{1}{R_j}}) & \text{if } \zeta_t \geq 1 \\ \zeta_t^{-1/\underline{R}}(\sum_{j=1}^{J}(\lambda_j e^{\rho_j t})^{-\frac{1}{R_j}}) & \text{if } \zeta_t < 1 \end{cases}$$

where $\overline{R} = \max_j R_j$ and $\underline{R} = \min_j R_j$. Hence

$$\zeta_t \leq \begin{cases} \Delta_t^{-\overline{R}}(\sum_{j=1}^{J}(\lambda_j e^{\rho_j t})^{-1/R_j})^{\overline{R}} & \text{if } \zeta_t \geq 1 \\ \Delta_t^{-\underline{R}}(\sum_{j=1}^{J}(\lambda_j e^{\rho_j t})^{-1/R_j})^{\underline{R}} & \text{if } \zeta_t < 1. \end{cases} \qquad (7.21)$$

Next, if we define $\kappa = \min_j \frac{\rho_j}{R_j}$ then

$$(\sum_{j=1}^{J}(\lambda_j e^{\rho_j t})^{-1/R_j})^{\overline{R}} \leq C_1 e^{-\kappa t\overline{R}} \qquad (7.22)$$

$$(\sum_{j=1}^{J}(\lambda_j e^{\rho_j t})^{-1/R_j})^{\underline{R}} \leq C_2 e^{-\kappa t\underline{R}} \qquad (7.23)$$

for some constants C_1, C_2. Therefore using (7.21), (7.22) and (7.23) we deduce that for convergence it is sufficient to find conditions such that

$$\int_0^\infty \mathbb{E}_t[\zeta_s \Delta_s] ds \le \int_0^\infty \mathbb{E}_t[\Delta_s^{1-\theta} e^{-\kappa s \theta}] ds < \infty$$

for $\theta = \underline{R}, \overline{R}$. Denoting $X_i(t) = \sigma_i W_t^i + \mu_i t$, we see that

$$\begin{aligned}\int_0^\infty \mathbb{E}_t[\Delta_s^{1-\theta} e^{-\kappa s \theta}] ds &\le C_3 \int_0^\infty e^{-\kappa s \theta} \mathbb{E}_t[\sum_i e^{(1-\theta)X_i(s)}] ds \\ &= \sum_i \int_0^\infty \exp[-\kappa \theta s + \frac{1}{2}\sigma_i^2 s(1-\theta)^2 + \mu_i s(1-\theta)] ds.\end{aligned}$$

where C_3 is a constant. So for convergence it is sufficient to insist each i that

$$\mu_i(1-\theta) + (1-\theta)^2 \frac{\sigma_i^2}{2} - \kappa\theta < 0 \tag{7.24}$$

7.4 Calculating the Interest Rate

As we have already mentioned, the interest rate is endogenously determined. It is useful to calculate the interest rate before and after a shock to the equilibrium and investigate how this rate is affected by the shocks we impose. In this section we derive an explicit expression for the interest rate in our modelling framework. To do this, observe that if we differentiate

$$\begin{aligned}\zeta_t &\equiv \exp(-\int_0^t r_s ds). \left.\frac{d\mathbb{P}^*}{d\mathbb{P}}\right|_{\mathcal{F}_t} \\ &\equiv \exp(-\int_0^t r_s ds) Z_t\end{aligned} \tag{7.25}$$

we get

$$\begin{aligned}d\zeta_t &= -r_t \zeta_t dt + \exp(-\int_0^t r_s ds) dZ_t \\ &= -r_t \zeta_t dt + \frac{\zeta_t}{Z_t} dZ_t \\ &= \zeta_t (Z_t^{-1} dZ_t - r_t dt),\end{aligned} \tag{7.26}$$

and so we see that the finite variation term in $\frac{d\zeta_t}{\zeta_t}$ is $-r_t dt$. It is this fact that we shall use to derive our expression.

First recall that the market clearing condition (7.14) says

$$\Delta_t = \sum_{j=1}^{J} (\zeta_t \lambda_j e^{\rho_j t})^{-\frac{1}{R_j}}.$$

This implies

$$\zeta_t = \phi(t, \Delta_t) \tag{7.27}$$

for some function ϕ and if we set $\psi = \phi^{-1}$ we have that

$$\psi(t, \phi(t, \Delta_t)) = \Delta_t. \tag{7.28}$$

From (7.27)[3]

$$d\zeta_t = \left(\frac{\partial \phi}{\partial t}\right) dt + \left(\frac{\partial \phi}{\partial \Delta_t}\right) d\Delta_t + \frac{1}{2}\left(\frac{\partial^2 \phi}{\partial \Delta_t^2}\right) d < \Delta >_t . \tag{7.29}$$

Next, our specification that the dividend stream, $\delta_i(t)$, is log Brownian motion for all i gives us that

$$\begin{aligned} d\Delta_t &= d\left(\sum_{i=1}^{M} \delta_i(t)\right) \\ &= \sum_{i=1}^{M} d(\delta_i(t)) \\ &= \sum_{i=1}^{M} \delta_i(t) \left[\sigma_i dW_t^i + \mu_i dt + \frac{1}{2}\sigma_i^2 dt\right]. \end{aligned} \tag{7.30}$$

Therefore, (7.29) together with (7.30) give

$$\begin{aligned} d\zeta_t = \left[\left(\frac{\partial \phi}{\partial t}\right) + \left(\frac{\partial \phi}{\partial \Delta_t}\right) \sum_{i=1}^{M} \delta_i(t) \{\mu_i + \sigma_i^2\} + \frac{1}{2}\left(\frac{\partial^2 \phi}{\partial \Delta_t^2}\right) \sum_{i=1}^{M} \delta_i^2(t)\sigma_i^2\right] dt \\ + \left(\frac{\partial \phi}{\partial \Delta_t}\right) \sum_{i=1}^{M} \delta_i(t)\sigma_i dW_t^i. \end{aligned} \tag{7.31}$$

What remains is to find expressions for the differentials

[3] We have omitted including the specific parameter dependencies for each function to ease reading.

$$\frac{\partial \phi}{\partial t}, \frac{\partial \phi}{\partial \Delta_t} \text{ and } \frac{\partial^2 \phi}{\partial \Delta_t^2}.$$

Unfortunately, because we cannot explicitly write down ϕ, calculating the partial derivatives above becomes more complicated and we must proceed by differentiating (7.28); we have

$$\frac{\partial \psi}{\partial t} + \frac{\partial \psi}{\partial \phi}\frac{\partial \phi}{\partial t} = 0 \tag{7.32}$$

and

$$\frac{\partial \psi}{\partial \phi}\frac{\partial \phi}{\partial \Delta_t} = 1. \tag{7.33}$$

Then differentiate (7.33) to give

$$\frac{\partial^2 \psi}{\partial \phi^2}\left(\frac{\partial \phi}{\partial \Delta_t}\right)^2 + \frac{\partial \psi}{\partial \phi}\frac{\partial^2 \phi}{\partial \Delta_t^2} = 0. \tag{7.34}$$

We rearrange (7.32), (7.33) and (7.34) so that

$$\begin{aligned}
\frac{\partial \phi}{\partial t} &= -\left(\frac{\partial \psi}{\partial t}\right) / \left(\frac{\partial \psi}{\partial \phi}\right) &(7.35)\\
\frac{\partial \phi}{\partial \Delta_t} &= \frac{\partial \phi}{\partial \psi} &(7.36)\\
\frac{\partial^2 \phi}{\partial \Delta_t^2} &= -\left(\frac{\partial^2 \psi}{\partial \phi^2}\right) / \left(\frac{\partial \psi}{\partial \phi}\right)^3. &(7.37)
\end{aligned}$$

In this manner, we have altered the problem into one of finding

$$\frac{\partial \psi}{\partial t}, \frac{\partial \psi}{\partial \phi} \text{ and } \frac{\partial^2 \psi}{\partial \phi^2},$$

which is more straightforward to handle. Using

$$\psi(t, \phi(t, \Delta_t)) = \psi(t, \zeta_t) = \sum_{j=1}^{J} \left(\zeta_t \lambda_j e^{\rho_j t}\right)^{-\frac{1}{R_j}}$$

we have

$$\frac{\partial \psi}{\partial t} = \sum_{j=1}^{J}\left(-\frac{\rho_j}{R_j}\right)\left(\zeta_t\lambda_j e^{\rho_j t}\right)^{-\frac{1}{R_j}} \tag{7.38}$$

$$\frac{\partial \psi}{\partial \phi} = \sum_{j=1}^{J}\left(-\frac{1}{R_j\zeta_t}\right)\left(\zeta_t\lambda_j e^{\rho_j t}\right)^{-\frac{1}{R_j}} \tag{7.39}$$

$$\frac{\partial^2 \psi}{\partial \phi^2} = \sum_{j=1}^{J}\left(\frac{1}{R_j\zeta_t}\right)^2\left(\zeta_t\lambda_j e^{\rho_j t}\right)^{-\frac{1}{R_j}} + \sum_{j=1}^{J}\left(\frac{1}{R_j\zeta_t^2}\right)\left(\zeta_t\lambda_j e^{\rho_j t}\right)^{-\frac{1}{R_j}}. \tag{7.40}$$

Now since the finite variation term of $\frac{d\zeta_t}{\zeta_t}$ is $-r_t dt$, we can put together (7.35), (7.36), (7.37), (7.38), (7.39), (7.40) and by using (7.31) we have that

$$r_t = \frac{\sum_{j=1}^{J}\left(\frac{\rho_j}{R_j}\right)\left(\zeta_t\lambda_j e^{\rho_j t}\right)^{-\frac{1}{R_j}}}{\sum_{j=1}^{J}\left(\frac{1}{R_j}\right)\left(\zeta_t\lambda_j e^{\rho_j t}\right)^{-\frac{1}{R_j}}} + \frac{\sum_{i=1}^{M}\delta_i(t)\left\{\mu_i + \frac{1}{2}\sigma_i^2\right\}}{\sum_{j=1}^{J}\left(\frac{1}{R_j}\right)\left(\zeta_t\lambda_j e^{\rho_j t}\right)^{-\frac{1}{R_j}}}$$
$$-\frac{1}{2}\left(\frac{\sum_{j=1}^{J}\left(\frac{1}{R_j}\right)^2\left(\zeta_t\lambda_j e^{\rho_j t}\right)^{-\frac{1}{R_j}} + \sum_{j=1}^{J}\left(\frac{1}{R_j}\right)\left(\zeta_t\lambda_j e^{\rho_j t}\right)^{-\frac{1}{R_j}}}{\left(\sum_{j=1}^{J}\left(\frac{1}{R_j}\right)\left(\zeta_t\lambda_j e^{\rho_j t}\right)^{-\frac{1}{R_j}}\right)^3}\right)\left(\sum_{i=1}^{M}\delta_i^2(t)\sigma_i^2\right) \tag{7.41}$$

This is the simplest form we can find for the interest rate in this situation.

7.5 Remarks

The work presented in this chapter is as far as we can proceed analytically in this general case. If we were to impose further assumptions, such as that R_j was the same for all j, or that $R_j = 1$ we could get closer to an explicit form for some of the expectations. However, in all cases we will eventually have to consider numerical methods.

Chapter 8

Numerical Examples: procedure, results and analysis

The closing comments of Chapter 7 suggested that it was not possible to get any further with an analytical solution and that numerical methods were necessary. Therefore in this chapter we will consider two explicit examples numerically; the first is the case where all agents have the utility function,

$$U_j(t,c) = e^{-\rho_j t}\frac{c^{1-R}}{(1-R)}, \tag{8.1}$$

so essentially the R_j's (CRRA coefficient) are the same for all agents. The second example uses the original utility function given in (7.2) taking the impatience parameter ρ_j to be the same for all agents.

One of the advantages of considering the first case where all agents have the same R_j is that it allows us to proceed a little further analytically and consequently simplifies the work we must do numerically. It will also be easier to analyse the results and carry out some comparative statics for this case.

The plan for this chapter is to begin in §8.1 by presenting the method we use to find the equilibrium and to describe how we synthesize a shock to the risky assets. The procedure we give can be looked at as being the key steps of the numerical code. In §8.2 we discuss the numerical differentiation, integration and root finding procedures used. §8.3 gives the analysis and equations we require for Example A, that is the case where all agents have the same risk aversion coefficient. The results obtained from the numerics and the discussion for this example are given in §8.4.

Next, §8.5 has the analysis for Example B where the investors can have different risk preferences but the same impatience value. The numerical results and discussion for this example are in §8.6. Finally, we end the chapter and this part of the thesis in §8.7 with some conclusions.

8.1 The Procedure

In this modelling procedure we begin by computing the initial equilibrium parameter values. This is under the assumption that the existing economic conditions prevail forever. Having done this, we then hit the equilibrium with a sudden shock, modifying the dynamics of one of the risky assets. To determine the reaction induced by the shock, it is necessary to recompute the equilibrium under the new dynamics and then examine the changes that have occurred.

8.1.1 Specific details

To reduce computation time, and as we are not concerned about the specific initial portfolios of the agents, we shall begin by specifying values for λ_j and then use these to find the corresponding equilibrium portfolio holdings for each agent. The specific steps in creating our two-stage (continuous framework) static equilibrium model are as follows.

1. We start by choosing appropriate values for λ_j, the Lagrange multiplier. These can either be chosen randomly, or from empirical experience and the specific values are unimportant provided each λ_j is unique and they are all positive. Next, observe that the share price of any asset i and the wealth of any agent j is given by

$$S_i(t) = f_i(t, \delta(t); \lambda) \text{ and } w_j(t) = g_j(t, \delta(t); \lambda)$$

   for some functions f_i and g_j which we can typically only determine numerically.

2. Using our numerical values for the asset prices and wealths of the agents, we can determine the initial portfolio, $\pi_j(0)$, and the initial amount held in the riskless deposit account, $\phi_j(0)$, that agent j would need to start from in order to get the desired consumption stream $c_j(.)$.

   For this, we observe that by applying Itô's formula to the above we obtain

$$dw_j(t) = \sum_{p=1}^{M} \frac{\partial g_j}{\partial \delta_p(t)} d\delta_p(t) + \text{ finite variation terms,}$$

and similarly from (7.8) we have

$$dS^i(t) = \sum_{p=1}^{M} \frac{\partial f_i}{\partial \delta_p(t)} d\delta_p(t) + \text{ finite variation terms.}$$

Next, using the change in wealth equation (7.5) and equating quadratic variation terms only, we get

$$\sum_{i=1}^{M} \pi_j^i(t) \frac{\partial f_i}{\partial \delta_p(t)} = \frac{\partial g_j}{\partial \delta_p(t)} \quad j = 1, .., J, \quad p = 1, .., M. \tag{8.2}$$

Therefore, by numerically finding

$$\frac{\partial f_i(0, \delta(0); \lambda)}{\partial \delta_p} \text{ and } \frac{\partial g_j(0, \delta(0); \lambda)}{\partial \delta_p},$$

we can determine $\pi_j(0)$ using the relationship (8.2). Moreover, since we know $w_j(0)$ $(= g_j(0, \delta(0); \lambda))$ and $\sum_{i=1}^{M} \pi_j^i(0) S_i(0)$, we can also determine the new holdings in the riskless asset, $\phi(0)$.

We now have the time-0 initial equilibrium values for the wealth and portfolio holdings of each agent, as well as the share prices. It is instructional to also evaluate the payoff $\mathbb{E}_0 [\int_0^\infty U_j(s, c_j(s))ds]$ for each agent $j = 1, .., J$.

3. At this point we hit the market with some kind of shock. This will be imposed via the dynamics of the risky assets, the details of which we discuss below.

4. Now using $\pi_j(0)$, $\phi_j(0)$ and our new post-shock (time 0+) dynamics for the assets, we compute $\tilde{\lambda}$ which solves

$$\sum_{k=1}^{M} \pi_j^k(0) f_k(0, \delta(0+); \tilde{\lambda}) + \phi_j(0) = g_j(0, \delta(0+); \tilde{\lambda}), \ j = 1, .., J. \tag{8.3}$$

5. Finally, armed with our new $\tilde{\lambda}$ we compute the new portfolio of risky assets and holdings in the bank ($\pi_j(0+)$ and $\phi_j(0+)$ $j = 1, .., J$) and examine what affect the shock has had on these. We also compute the new wealth and payoff for each agent and compare these to the values before the shock. The interest rate before and after the shock should also be calculated.

8.1.2 Synthesizing a shock to the market equilibrium

We have chosen to impose shocks to the market via the risky assets. Although there are a number of ways we could attempt to do this, we have decided to limit ourselves to imposing the shocks only via the dividend stream of the asset and to just hitting one asset at a time.

The shock can be imposed in one of two ways:

1. by hitting the mean return value, μ, of the dividend stream of one of the assets;
2. by hitting the initial dividend, $\delta(0)$ of an asset.

Therefore, under our construction, if we wished to synthesize a shock to asset i, we could either modify μ_i or $\delta_i(0)$ in the specification of $\delta_i(t)$. We will consider what happens if these parameters are suddenly increased or reduced by a fixed percentage.

8.2 Numerical considerations

It is already clear from the analysis in Chapter 7 and the modelling procedure given in §8.1, that to proceed numerically we must use a routine which can compute multi-dimensional integrals. Moreover, we also require numerical differentiation, minimisation and root finding routines; the uses for each of these will become apparent as we proceed.

The integration procedure which we use here is the routine D01FCF, supplied by NAG. This routine is capable of evaluating any multi-dimensional integral, up to 15 dimensions. This routine is ideal for our purposes as our numerical examples do not look at models with more than 4 assets (i.e. 4 dimensions).

This routine is used to compute expectations such as

$$\mathbb{E}_t[\zeta_s \delta_i(s)] \text{ and } \mathbb{E}_t[\Delta_u^{-R} \delta_i(u)].$$

The transition density of $\log \delta_i(t)$ is

$$p_i(t,x,y) = \frac{1}{\sqrt{2\pi\sigma_i^2 t}} \exp(-\frac{(y-x-\mu_i t)^2}{2t\sigma_i^2}) \tag{8.4}$$

where σ_i is the volatility and μ_i is the mean rate of return of asset i.

We also need to evaluate an integral over time (see (8.13) for example), this we do using the NAG routine D01AMF, a 1-D quadrature routine which allows for semi-infinite intervals.

The minimisation routine which we will use is E04JYF, this gives us a tool to solve the set of J equations when we compute $\tilde{\lambda}$ after the shock[1].

The numerical differentiation of $S_i(t) = f_i(t, \delta(t); \lambda)$ and $w_j(t) = g_j(t, \delta(t); \lambda)$ is standard, we use the $O(h^2)$ numerical approximation of the derivative of a function $p(x)$ at the point x_0

$$\frac{p(x_0 + h) - p(x_0 - h)}{2h}, \tag{8.5}$$

this is the 2[nd] order central difference formula.

Finally, we also ask the numerical code to produce a couple of extra diagnostic parameters whose values we shall compare before and after the shock to aid the interpretation of the output.

The first diagnostic value we compute is

$$\nu_j = \left(\frac{V_j(0+)}{V_j(0)}\right)^{\frac{1}{1-R_j}}, \tag{8.6}$$

for all j, where $V_j(0)$ is the payoff of agent j before the shock and $V_j(0+)$ is the payoff after the shock.

This quantity may be looked at as a true measure of how much better or worse things have got after the shock. If agent j's original consumption stream was multiplied by the constant ν_j, then the original payoff $V_j(0)$ would be changed to the post-shock payoff $V_j(0+)$.

The second value we shall consider is the changeover time. All our runs will be done with two risky assets; to begin with, the dividend rate of one of the assets will be longer than the other, however, there will be a point in time when the second asset becomes preferred. We refer to this as the changeover time. It is computed by finding the time

[1] Recall that this is the same routine that was used in our work on calibrating potential models in Chapter 6

t when

$$\mathbb{E}[\delta_0(t) - \delta_1(t)] = 0. \tag{8.7}$$

8.3 Example A: all agents have the same CRRA

In this section we will consider the situation where all agents are equally risk averse. Therefore the utility for each j is

$$U_j(t, c_j(t)) = e^{-\rho_j t}\frac{c_j^{1-R}(t)}{1-R}, \tag{8.8}$$

and we assume that $R > 1$ throughout. One of the key advantages behind this choice of utility function is that it is now possible to explicitly write down the state price density process, ζ_t. We have from the market clearing condition that

$$\Delta_t = \sum_{j=1}^{J}(\zeta_t\lambda_j e^{\rho_j t})^{-1/R} \tag{8.9}$$

hence,

$$\zeta_t = \frac{[\sum_{j=1}^{J}(e^{\rho_j t}\lambda_j)^{-1/R}]^R}{\Delta_t^R}. \tag{8.10}$$

For ease of notation, we write $\psi(t) \equiv \sum_{j=1}^{J}(e^{\rho_j t}\lambda_j)^{-1/R}$.

Some slight simplifications of the expressions for the share price and wealth given in §7.3.1 are now possible. The wealth process of agent j is

$$\begin{aligned}
w_j(t) &= \frac{1}{\zeta_t}\mathbb{E}_t[\int_t^\infty \zeta_s c_j(s)ds] \\
&= \left[\frac{\Delta_t}{\psi(t)}\right]^R .\mathbb{E}_t[\int_t^\infty \zeta_s^{(1-1/R)}\lambda_j^{(-1/R)}e^{\frac{-\rho_j s}{R}}ds] \\
&= \left[\frac{\Delta_t}{\psi(t)}\right]^R .\mathbb{E}_t[\int_t^\infty \psi^{(R-1)}(s)\lambda_j^{(-1/R)}e^{\frac{-\rho_j s}{R}}\Delta_s^{1-R}ds] \\
&= \left[\frac{\Delta_t}{\psi(t)}\right]^R \int_t^\infty \psi^{(R-1)}(s)\lambda_j^{(-1/R)}e^{\frac{-\rho_j s}{R}}\mathbb{E}_t[\Delta_s^{1-R}]ds \\
&= \left[\frac{\Delta_t}{\psi(t)}\right]^R \int_t^\infty (\sum_{k=1}^{J}(e^{\rho_k s}\lambda_k)^{-1/R})^{R-1}\lambda_j^{-1/R}e^{\frac{-\rho_j s}{R}}\mathbb{E}_t[\Delta_s^{1-R}]ds.
\end{aligned} \tag{8.11}$$

Similarly, for the price of share i we have

$$\begin{aligned} S_i(t) &= \frac{1}{\zeta_t}\mathbb{E}_t[\int_t^\infty \zeta_u \delta_i(u)du] \\ &= \left[\frac{\Delta_t}{\psi(t)}\right]^R .\mathbb{E}_t[\int_t^\infty \psi^R(u)\Delta_u^{-R}\delta_i(u)du] \\ &= \left[\frac{\Delta_t}{\psi(t)}\right]^R \int_t^\infty (\sum_{k=1}^J (e^{\rho_k u}\lambda_k)^{-1/R})^R \mathbb{E}_t[\Delta_u^{-R}\delta_i(u)]du. \end{aligned} \tag{8.12}$$

It follows that the J equations we must solve to compute $\hat{\lambda}$ in the post shock equilibrium are given by (using the wealth equation as described in §7.3.3)

$$\left[\frac{\Delta_0}{\psi(0)}\right]^R \int_0^\infty \left[(\sum_{k=1}^J (e^{\rho_k u}\lambda_k)^{-1/R})^R \mathbb{E}_0[\Delta_u^{-R}\sum_{i=1}^M \pi_j^i(0)\delta_i(u)] \right. \\ \left. -(\sum_{k=1}^J (e^{\rho_k u}\lambda_k)^{-1/R})^{R-1}\lambda_j^{-1/R}e^{\frac{-\rho_j u}{R}}\mathbb{E}_0[\Delta_u^{1-R}]\right] du + \phi_j(0) = 0 \; j = 1,..,J. \tag{8.13}$$

We can also express the payoff of agent j as

$$V_j(0) \equiv \frac{\lambda_j}{1-R}\left(\frac{\psi_0}{\Delta_0}\right)^R w_j(0) \tag{8.14}$$

The sufficient condition for convergence of (8.13) is simpler than for the case of different R_j presented in §7.3.4.

Assuming $R > 1$, $\rho_1 = \min_j \rho_j$ and denoting $X_i(t) = \sigma_i W_t^i + \mu_i t$, we see that

$$\begin{aligned} \int_0^\infty e^{-\rho_1 u}\mathbb{E}[\Delta_u^{1-R}]du &\leq \int_T^\infty e^{-\rho_1 t}\mathbb{E}[\sum_i e^{(1-R)X_i(t)}]dt \\ &= \sum_i \int_T^\infty \exp[-\rho_1 t + (1-R)^2\frac{1}{2}\sigma_i^2 t + \mu_i(1-R)t]dt. \end{aligned}$$

So the condition that for each i

$$\mu_i(1-R) + (1-R)^2\frac{\sigma_i^2}{2} - \rho_1 < 0 \tag{8.15}$$

guarantees convergence of (8.13).

Also, by taking $R_j \equiv R$ for all j in (7.41) we get the following, simpler form, for the interest rate in this example,

$$r_t = \left(\frac{1}{\psi(t)}\right)\left\{\sum_{j=1}^{J}\rho_j(\lambda_j e^{\rho_j t})^{-\frac{1}{R}}\right\} + \left(\frac{R}{\Delta_t}\right)\sum_{i=1}^{M}\delta_i(t)\left\{\mu_i + \frac{1}{2}\sigma_i^2\right\}$$

$$- \frac{R(R+1)}{2\Delta_t^2}\left\{\sum_{i=1}^{M}\delta_i^2(t)\sigma_i^2\right\}. \quad (8.16)$$

We will often about (8.16) as being made up of three terms.

8.4 Results and Analysis for Example A

Our market has two risky assets and four agents.

Asset 0 is the more volatile asset. The initial dividend stream is low when compared to the other asset but the mean return and the volatility are significantly higher than that of asset 1.

Asset 1 has been set up to act as a more stable and secure asset. Although the mean return is lower than that of the other asset, the volatility is significantly lower as well. The initial dividend stream, δ_0 is high for this asset.

The parameter values for the two assets are given in the Table 8.1.

Asset	μ	σ	$\delta(0)$	$\mu - \frac{1}{2}\sigma^2$
0	0.26	0.38	2.15	0.1878
1	0.14	0.342	17.6	0.081518

Table 8.1: Example A: initial parameter values for the two risky assets in the model

We could interpret the assets as follows: Asset 0 is the stock index in an Asian economy, with higher mean return and volatility than Asset 1, which represents a larger more stable economy.

The parameter settings together with a description of the kind of market participants that might be described by these parameters are given in Table 8.2.

Market Participants					
Agent	ρ	R (CRRA)	Possible Characteristics		λ
			Type of Business	Market Outlook	
0	0.05	3.25	Pension Fund	Long Term	0.1
1	0.075	3.25	Large Multi-national	Mid/Long Term	4.0
2	0.09	3.25	Technology/Computers	Mid Term	22.0
3	0.12	3.25	Property Leasing Company	Short Term	29.0

Table 8.2: Example A: the characteristics of the four agents in the model

The first stage market equilibrium was found before we hit the market with a shock. The equilibrium portfolio for this market is given in Table 8.4.

Pre-Shock Market Equilibrium Values					
	Agent Holdings				Price (per unit)
	0	1	2	3	
Asset 0	0.589851 (21.74%)	0.191240 (22.48%)	0.113689 (22.92%)	0.105210 (23.76%)	16.091924
Asset 1	0.606244 (78.26%)	0.188247 (77.52%)	0.109166 (77.08%)	0.096346 (76.24%)	56.370393
Bank ϕ	0.0	0.0	0.0	0.0	
Wealth %	60.2603	18.8911	11.0171	9.8314	
Consumption $c_j(0)$	11.711705	3.764251	2.227792	2.046252	
Interest Rate	10.5603%				
Changeover Time	19.78 yrs				

Table 8.3: Example A: pre-shock market equilibrium parameter values

Notice that the amount each agent invests in the riskless asset is zero in the pre-shock equilibrium in Table 8.4 (and in all the equilibrium tables in Example A). This is a consequence (see Heritage & Rogers [55]) of the fact that all agents have the same coefficient of relative risk aversion. The figures for wealth % report the percentage of total market value held by each agent. Agent 0 is the most patient agent and we see from his portfolio that he prefers to hold more of Asset 1, whereas the other agents take the opposite position.

8.4.1 Hitting the initial dividend stream, $\delta(0)$

The first situation considered is imposing a 20% drop to the initial dividend stream $\delta(0)$ for each asset (taken one at a time). The post-shock equilibrium values for these two shocks are given in Tables 8.4 and 8.5.

Post-Shock Equilibrium (20% drop to $\delta(0)$ of Asset 0)					
	Agent Holdings				Price
	0	1	2	3	(per unit)
Asset 0	0.587610 (-)	0.191655 (+)	0.114319 (+)	0.106408 (+)	15.105926 (-)
Asset 1	0.606805 (+)	0.188138 (-)	0.109017 (-)	0.096042 (-)	59.911840 (+)
Bank ϕ	0.0 (NC)	0.0 (NC)	0.0 (NC)	0.0 (NC)	
Wealth %	60.2940 (+)	18.8846 (-)	11.0085 (-)	9.8129 (-)	
Consumption $c_j(0)$	11.456109 (-)	3.682329 (-)	2.179597 (-)	2.001965 (-)	
λ	0.100017 (+)	3.999865 (-)	21.989782 (-)	28.987446 (-)	
ν	0.953587	0.954056	0.954422	0.954879	
Interest Rate	7.2027% (-)				
Changeover Time	21.88 yrs (+)				

Table 8.4: Example A: post-shock equilibrium values - 20% drop to $\delta(0)$ of Asset 0

Comparing the pre-shock portfolios of each agent and their percentage holdings of the risky assets we might expect that if Asset 0 becomes less attractive, because of a reduction in $\delta_0(0)$ (or even of μ_0), then Agent 0 will improve his situation relative to the others. This is because he is less reliant on Asset 0. Conversely, if Asset 1 becomes less attractive, then we expect the reverse to happen. This behaviour is exactly what we observe in the results of Tables 8.4 and 8.5.

Another change that is predictable is that of the changeover time; by just considering how it is computed using (8.7) it is possible to determine the direction of change before imposing the shock.

Post-Shock Equilibrium (20% drop to $\delta(0)$ of Asset 1)					
	Agent Holdings				Price
	0	1	2	3	(per unit)
Asset 0	0.591731 (+)	0.190893 (-)	0.113172 (-)	0.104203 (-)	13.789734 (-)
Asset 1	0.605631 (-)	0.188361 (+)	0.109336 (+)	0.096672 (+)	42.471618 (-)
Bank ϕ	0.0	0.0	0.0	0.0	
Wealth %	60.2224 (-)	18.8981 (+)	11.0276 (+)	9.8518 (+)	
Consumption $c_j(0)$	9.623834 (-)	3.093462 (-)	1.830896 (-)	1.681808 (-)	
λ	0.100018 (+)	3.999596 (-)	21.994015 (-)	28.986028 (-)	
ν	0.842939	0.842671	0.842515	0.842180	
Interest Rate	14.3724% (+)				
Changeover Time	17.68 yrs (-)				

Table 8.5: Example A: post-shock equilibrium values - 20 % drop to $\delta(0)$ of Asset 1

8.4.2 Hitting the mean rate of return, μ

Next we impose shocks to the mean rate of return. The first case is hitting μ_0 by a 20% drop, see Table 8.6, and the second case is identical but for Asset 1 (Table 8.7).

The rationale for the behaviour observed when we imposed a shock to the initial dividend stream in §8.4.1 applies equally here when the mean rate of return value is hit. Something which is much harder to justify is the effect on ν in all the post-shock tables in Example A; we consistently see that the agent whose change in percentage wealth is best actually fares worst in terms of the effect on the payoff ν value.

The movements of the (time-0) interest rate after a shock has been imposed are also very hard to predict, by looking at (8.16) it is immediately obvious that when we drop the initial dividend stream and recompute the equilibrium, it has an effect on all three terms in the expression. When dropping the mean rate of return the effect is limited to the first two terms but even here we can't predict what will happen to the first term and so we're stuck. Numerical tests have indicated that the change in the first term, which is actually just a weighted average of ρ_j, is much smaller than the change in the second term, but it is not possible to verify that this would always be the case.

Post-Shock Equilibrium (20% drop to μ of Asset 0)					
	Agent Holdings				Price (per unit)
	0	1	2	3	
Asset 0	0.588929 (-)	0.191430 (+)	0.113953 (+)	0.105689 (+)	17.463968 (+)
Asset 1	0.606495 (+)	0.188195 (-)	0.109094 (-)	0.096215 (-)	64.035833 (+)
Bank ϕ	0.0	0.0	0.0	0.0	
Wealth %	60.2731 (+)	18.8888 (-)	11.0135 (-)	9.8245 (-)	
Consumption $c_j(0)$	11.685055 (-)	3.768967 (+)	2.235098 (+)	2.060881 (+)	
λ	0.100630 (+)	3.979301 (-)	21.742807 (-)	28.304617 (-)	
ν	0.945897	0.950873	0.953738	0.959222	
Interest Rate	8.7278% (-)				
Changeover Time	38.731643 yrs (+)				

Table 8.6: Example A: post-shock equilibrium values - 20% shock to μ of Asset 0

Post-Shock Equilibrium (20% drop to μ of Asset 1)					
	Agent Holdings				Price (per unit)
	0	1	2	3	
Asset 0	0.584359 (-)	0.191220 (-)	0.114261 (+)	0.106734 (+)	20.355713 (+)
Asset 1	0.608000 (+)	0.188255 (+)	0.108985 (-)	0.095857 (-)	63.632045 (+)
Bank ϕ	0.0	0.0	0.0	0.0	
Wealth %	60.2270 (-)	18.8974 (+)	11.0263 (+)	9.8493 (+)	
Consumption $c_j(0)$	11.684024 (-)	3.769037 (+)	2.235349 (+)	2.061591 (+)	
λ	0.100628 (+)	3.977819 (-)	21.728098 (-)	28.264127 (-)	
ν	0.933538	0.938086	0.940743	0.945896	
Interest Rate	2.4585% (-)				
Changeover Time	15.656835 yrs (-)				

Table 8.7: Example A: post-shock equilibrium values - 20% shock to μ of Asset 1

8.5 Explicit Example B: agents have different CRRA

In this case we consider the situation where the agents each have different risk preferences but the same impatience preference ρ. Therefore agent j has utility

$$U_j(t,c) = e^{-\rho t}\frac{c^{1-R_j}}{(1-R_j)},$$

and the market clearing condition in this example is

$$\Delta_t \equiv \sum_{i=1}^{M}\delta_i(t) = \sum_{j=1}^{J}c_j(t) = \sum_{j=1}^{J}(\zeta_t\lambda_j e^{\rho t})^{-\frac{1}{R_j}}. \tag{8.17}$$

Unfortunately, unlike in Example A, we cannot solve (8.17) explicitly for ζ_t in terms of Δ_t; indeed the best we can do is to determine ζ_t numerically using a root finder but this inevitably slows the computations down.

The algorithm we will use to find ζ_t numerically is the Newton-Raphson approximation; this is not computationally intensive and should converge quickly provided we give it a suitable starting approximation. Therefore, we seek a tight lower bound for ζ_t to start the iterative process. Define $\overline{R} = \max_j R_j$ and $\underline{R} = \min_j R_j$, then if $\zeta_t < 1$ we have

$$\begin{aligned}
\Delta_t &= \sum_{j=1}^{J}(\zeta_t\lambda_j e^{\rho_j t})^{-\frac{1}{R_j}} \\
&= \sum_{j=1}^{J}\zeta_t^{-\frac{1}{R_j}}(\lambda_j e^{\rho_j t})^{-\frac{1}{R_j}} \\
&\geq \sum_{j=1}^{J}\zeta_t^{-1/\overline{R}}(\lambda_j e^{\rho_j t})^{-\frac{1}{R_j}}.
\end{aligned}$$

Therefore,

$$\zeta_t \geq \left(\frac{\sum_{j=1}^{J}(\lambda_j e^{\rho_j t})^{-\frac{1}{R_j}}}{\Delta_t}\right)^{\overline{R}}, \tag{8.18}$$

and similarly for $\zeta_t \geq 1$

$$\zeta_t \geq \left(\frac{\sum_{j=1}^{J} (\lambda_j e^{\rho_j t})^{-\frac{1}{R_j}}}{\Delta_t} \right)^{\underline{R}} . \tag{8.19}$$

The expressions for the wealth of each agent and the share prices are the same as those given in the analysis in §7.3.1 except that $\rho_j = \rho$. Indeed, the other expressions derived in Chapter 7: the equations governing an equilibrium solution, the convergence condition, the interest rate and the payoff expression translate directly in the same way.

8.6 Results and Analysis for Example B

We again have two assets and four agents. The dynamics of the two assets have similar characteristics to the assets in Example A (see §8.4).

The dynamics of the assets in this example are given in Table 8.8.

Asset	μ	σ	$\delta(0)$	$\mu - \frac{1}{2}\sigma^2$
0	0.245	0.426	3.0	0.154262
1	0.125	0.34	20.0	0.067200

Table 8.8: Example B: initial parameter values for the two risky assets in the model

Again we have developed the four agents in the market so that they have their own individual characteristics and market outlook; these obviously relate to the chosen CRRA value for each agent. We take CRRA values between 2 and 9. Details of the market participants are given in Table 8.9.

Market Participants					
Agent	ρ	R (CRRA)	Possible Characteristics		λ
			Type of Business	Market Outlook	
0	0.04	2.5	Currency Speculator	Adventurous	8.0
1	0.04	4.5	Computer manufacturer	Entrepreneurial	14.0
2	0.04	6.5	Chemical engineer	Steady	20.0
3	0.04	8.5	Pension fund	Cautious	3.7

Table 8.9: Example B: the characteristics of the four agents in the model

The first stage, pre-shock, market equilibrium values are given in Table 8.10

Pre-Shock Market Equilibrium Values					
	Agent Holdings				Price (per unit)
	0	1	2	3	
Asset 0	1.455557 (60.21%)	-0.056750 (-8.61%)	-0.162100 (-38.07%)	-0.236682 (-54.65%)	26.999982
Asset 1	0.605079 (73.47%)	0.168260 (74.94%)	0.111374 (76.79%)	0.115294 (78.14%)	79.250010
Bank ϕ	-21.982811 (-33.68%)	5.991206 (33.67%)	7.044518 (61.29%)	8.945703 (76.51%)	
Wealth % %	61.4351	16.7456	10.8166	11.0027	
Consumption $c_j(0)$	14.314344	3.873350	2.417163	2.395143	
Interest Rate	6.5231%				
Changeover Time	21.79 yrs				

Table 8.10: Example B: pre-shock market equilibrium parameter values

We see from Table 8.10 that Agent 0 invests heavily in Asset 0 which is the high-return, high-risk asset, whereas the other agents go short in this asset. This is what we would expect given that Agent 0 is 'adventurous' in his outlook; indeed, this agent actually shorts cash to take up this risky position. It is clear in this example that the portfolio of the 'adventurous' agent is tilted toward the riskier asset, while the other agents lean towards the less risky asset. Their attitude towards risk is also reflected in the different proportions of wealth invested in the bank deposit account; for Agent 1 it is 33.67%, Agent 2 has 61.29% and Agent 3 has 76.51% of wealth in the bank.

8.6.1 Hitting the initial dividend stream, $\delta(0)$

The two shocks we look at here are hits applied to the dividend stream $\delta(0)$ of Asset 0 and Asset 1. In both cases we drop the original initial dividend by 20%, the results are presented in Tables 8.11 and 8.12 for shocks to Asset 0 and Asset 1 respectively.

Post-Shock Equilibrium (20% drop to $\delta(0)$ of Asset 0)					
	Agent Holdings				Price
	0	1	2	3	(per unit)
Asset 0	1.630700 (+)	-0.093996 (-)	-0.219603 (-)	-0.316776 (-)	25.709929 (-)
Asset 1	0.580465 (-)	0.174189 (+)	0.119373 (+)	0.125947 (+)	84.653931 (+)
Bank ϕ	-24.410750 (-)	6.448339 (+)	7.848728 (+)	10.107583 (+)	
Wealth %	60.4004 (-)	17.0125 (+)	11.1503 (+)	11.4369 (+)	
Consumption $c_j(0)$	13.867642 (-)	3.794971 (-)	2.377321 (-)	2.360066 (-)	
λ	7.920497 (-)	14.038506 (+)	20.379257 (+)	3.836222 (+)	
ν	0.940726	0.958420	0.965115	0.968889	
Interest Rate	2.6932% (-)				
Changeover Time	24.35 yrs(+)				

Table 8.11: Example B: post-shock equilibrium values - 20% shock to $\delta(0)$ of Asset 0

Notice from the tables that agents who are willing to take less risks suffer least in terms of the reduction in their ν value. This seems logical when you consider the holdings of the various agents in the different assets, and also bear in mind that risk averse agents invest more in the riskless bank account.

Next, observe that the changes in the bank holdings and holdings of Asset 0 move in opposite directions when we hit Asset 0 and Asset 1, but the changes in the holdings of Asset 1 move in the same direction in both cases. It is much harder to formulate any hard rules or use comparative statics to explain this behaviour.

Post-Shock Equilibrium (20% drop to $\delta(0)$ of Asset 1)					
	Agent Holdings				Price
	0	1	2	3	(per unit)
Asset 0	1.410534 (-)	-0.025031 (+)	-0.150491 (+)	-0.235526 (+)	23.174498 (-)
Asset 1	0.567720 (-)	0.178114 (+)	0.123396 (+)	0.130819 (+)	60.097916 (-)
Bank ϕ	-18.697785 (+)	4.656619 (-)	6.057909 (-)	7.992218 (-)	
Wealth %	57.7781 (-)	17.7488 (+)	11.9907 (+)	12.4823 (+)	
Consumption $c_j(0)$	11.121531 (-)	3.405024 (-)	2.221998 (-)	2.251447 (-)	
λ	8.355157 (+)	13.893945 (-)	19.210199 (-)	3.478896 (-)	
ν	0.781772	0.895363	0.935604	0.954849	
Interest Rate	9.2766% (+)				
Changeover Time	19.23 yrs(-)				

Table 8.12: Example B: post-shock equilibrium values - 20% shock to $\delta(0)$ of Asset 1

8.6.2 Hitting the mean rate of return, μ

In this section we apply shocks to the market by changing the μ value. The first case is hitting μ of Asset 0 with a 20% drop, the results for which are shown in Table 8.13. We also hit the μ value of Asset 1 with a similar shock, see Table 8.14 for this case.

The numerical results when we hit the mean rate of return again suggest few rules we can rely on. We see that in both cases the moves of holdings in Asset 0 and the bank deposit account are in the same directions. The moves in Asset 1 are in opposite directions.

Notice that when we hit μ, Agent 0 ends up better off after both shocks, that is, his ν value is higher. Perhaps this is not so surprising in the case where μ_1 is hit because this agent took a large position in Asset 0 which ended up stronger after the shock, but if this is the correct rationale then the rise in ν when μ_0 is hit is unexpected! One explanation may be that the relative value of cash holdings in the bank fell after the shock to μ_0, so Agent 0's short position in cash helped overall.

Post-Shock Equilibrium (20% drop to μ of Asset 0)					
	Agent Holdings				Price
	0	1	2	3	(per unit)
Asset 0	1.493152 (+)	-0.073308 (-)	-0.173121 (-)	-0.246852 (-)	30.156079 (+)
Asset 1	0.640797 (+)	0.158548 (-)	0.099976 (-)	0.100696 (-)	93.146808 (+)
Bank ϕ	-26.439115 (-)	7.396607 (+)	8.437206 (+)	10.607758 (+)	
Wealth %	63.4872 (+)	16.1819 (-)	10.1598 (-)	10.1710 (-)	
Consumption $c_j(0)$	15.051483 (+)	3.675158 (-)	2.179820 (-)	2.093539 (-)	
λ	4.573532 (-)	11.493651 (-)	25.378463 (+)	7.528535 (+)	
ν	1.398462	0.957025	0.834503	0.773732	
Interest Rate	5.7322% (-)				
Changeover Time	49.84 yrs (+)				

Table 8.13: Example B: post-shock equilibrium values - 20% shock to μ of Asset 0

Post-Shock Equilibrium (20% drop to μ of Asset 1)					
	Agent Holdings				Price
	0	1	2	3	(per unit)
Asset 0	1.519370 (+)	-0.096709 (-)	-0.179252 (-)	-0.243504 (-)	33.723188 (+)
Asset 1	0.599342 (-)	0.170970 (+)	0.113061 (+)	0.116680 (+)	88.960809 (+)
Bank ϕ	-23.625790 (-)	7.097329 (+)	7.473290 (+)	9.053588 (+)	
Wealth %	65.9657 (+)	15.5243 (-)	9.3628 (-)	9.1472 (-)	
Consumption $c_j(0)$	15.521946 (+)	3.532356 (-)	2.030363 (-)	1.915336 (-)	
λ	2.479458 (-)	8.043255 (-)	23.577289 (+)	9.389464 (+)	
ν	2.16556	1.020627	0.79002	0.68943	
Interest Rate	1.7678% (-)				
Changeover Time	16.93 yrs (-)				

Table 8.14: Example B: post-shock equilibrium values - 20% shock to μ of Asset 1

8.7 Conclusions

In this part of the thesis we have taken a simple model for an economy where the structure of the equilibrium is well known, but for which it is very hard to say anything concrete about the equilibrium solution. Our objective has been to attempt to illuminate the behaviour of this type of equilibrium through applying sudden unexpected shocks to the economy.

As there appear to be no interesting tractable analytical cases we have pursued a numerical analysis of a few small examples. However, a market economy such as this, even in its simplified form with just 2 risky assets and 4 agents, is a difficult and intricate puzzle to attempt to unravel.

Shocks were applied to the dynamics of the dividend stream of a risky asset, which was modelled by a log Brownian process. Changes in the share price, the wealth of the agents, the holdings in the risky assets and riskless bank deposit and the endogenously determined interest rate were all recorded. Some of the numerical results could be justified to a limited extent, but in many cases it was impossible to find clear rules or comparative statics for particular movements. Moreover, because the equilibrium depended on everything, all agents and the characteristics of all assets, a change in any part of the system affected everything else, and so we cannot say that a given change was due solely to any one change elsewhere.

Appendix A

Calculation of extra terms in the Petrov expansion to the CLT

In this Appendix we will discuss how equations (2.9) and (2.10), obtained using Petrov's expansion to the CLT, were derived in Chapter 2.

We use the following theorem from Petrov [9].

Theorem A.1 (Theorem 5.22 Petrov [9]) *If we let $\{Y_n\}$ be a sequence of independent identically distributed random variables with mean $m(\Delta)$ and variance $\sigma(\Delta)^2$ then for $Z_n = Y_1 + Y_2 + .. + Y_n$ we have*

$$\mathbb{P}\left[\frac{Z_n - nm(\Delta)}{\sigma(\Delta)\sqrt{n}} \leq x\right] = \Phi(x) + \frac{\alpha(1-x^2)e^{-\frac{x^2}{2}}}{\sqrt{72\pi n}} + o(n^{-\frac{1}{2}})$$

where

$$\alpha = \mathbb{E}\left[\left(\frac{Y_1 - m(\Delta)}{\sigma(\Delta)}\right)^3\right].$$

Applying this theorem in the context of our problem in Chapter 2 we have

$$\mathbb{P}\left[\frac{\sum_{j=1}^{M}\zeta_j^* - M\varepsilon}{\nu\sqrt{M}} \leq x\right] = \Phi(x) + \frac{\alpha(1-x^2)e^{-\frac{x^2}{2}}}{\sqrt{72\pi M}} + o(M^{-\frac{1}{2}}) \tag{A.1}$$

where

$$\alpha = \mathbb{E}\left[\left(\frac{\zeta_1 - \varepsilon}{\nu}\right)^3 \middle| \zeta_1 \in [a,b]\right], \tag{A.2}$$

and $\zeta_j^* = (\zeta_j | \zeta_j \in [a,b])$ for $j = 1,..,M$.

Next, using the substitution $x = \frac{y-M\varepsilon}{\nu\sqrt{M}}$ and ignoring $o(M^{-1/2})$ terms in (A.1) yields

$$\mathbb{P}(\sum_{j=1}^{M} \zeta_j^* \leq y) = \Phi(\frac{y - M\varepsilon}{\nu\sqrt{M}}) + \frac{\alpha}{\sqrt{72\pi M}}\left(1 - \frac{(y-M\varepsilon)^2}{\nu^2 M}\right) e^{-\frac{(y-M\varepsilon)^2}{2M\nu^2}}. \tag{A.3}$$

Using the notation $F_M(y) = \mathbb{P}[\sum_{i=1}^{M} \zeta_i^* \leq y]$, observe that

$$\begin{aligned} \mathbb{E}\left[\left(K - S_0 \exp\left(\sum_{i=1}^{M} \zeta_i^*\right)\right)^+\right] &= \int_{-\infty}^{\log(\frac{K}{S_0})} (K - S_0 e^y) F_M(dy) \\ &= \int_{-\infty}^{\log(\frac{K}{S_0})} S_0 e^y F_M(y) dy. \end{aligned} \tag{A.4}$$

Therefore to find the extra term to add to our approximation (2.7), we find from (A.3) and (A.4) that the extra term is given by

$$\lambda \int_{-\infty}^{\log \frac{K}{S_0}} S_0 e^y \left(1 - \frac{(y - \varepsilon M)^2}{\nu^2 M}\right) e^{-\frac{(y-M\varepsilon)^2}{2\nu^2 M}} dy, \tag{A.5}$$

where $\lambda = \frac{\alpha}{\sqrt{72\pi M}}$. We first consider

$$A = \lambda S_0 \int_{-\infty}^{\log \frac{K}{S_0}} e^y \left(\frac{(y - M\varepsilon)^2}{\nu^2 M}\right) e^{-\frac{(y-M\varepsilon)^2}{2\nu^2 M}} dy. \tag{A.6}$$

It can be verified that

$$\begin{aligned} A &= K\lambda\left(M\varepsilon - \log\left(\frac{K}{S_0}\right)\right) e^{-\frac{(\log(K/S_0)-M\varepsilon)^2}{2\nu^2 M}} + S_0\lambda \int_{-\infty}^{\log \frac{K}{S_0}} (y - M\varepsilon) e^y e^{-\frac{(y-M\varepsilon)^2}{2\nu^2 M}} dy \\ &+ S_0\lambda \int_{-\infty}^{\log \frac{K}{S_0}} e^y e^{-\frac{(y-M\varepsilon)^2}{2\nu^2 M}} dy. \end{aligned} \tag{A.7}$$

Next, putting (A.7) into (A.5), because of cancelling, all that remains is to calculate

$$B = \lambda S_0 \int_{-\infty}^{\log \frac{K}{S_0}} e^y (y - M\varepsilon) e^{-\frac{(y-M\varepsilon)^2}{2\nu^2 M}} dy. \quad \text{(A.8)}$$

This integral is

$$B = -\lambda K M \nu^2 e^{-\frac{(\log(K/S_0)-M\varepsilon)^2}{2\nu^2 M}} + \lambda S_0 M \nu^2 (2\pi M \nu^2)^{1/2} e^{\frac{1}{2}M\nu^2 + M\varepsilon} \Phi(d_2), \quad \text{(A.9)}$$

where

$$d_2 = \frac{\log(\frac{K}{S_0}) - (M\nu^2 + M\varepsilon)}{\nu\sqrt{M}},$$

and $\Phi(.)$ is the cumulative distribution function with zero mean and unit standard deviation.

Although the integrals (A.6) and (A.8) were originally calculated by hand, the calculus involved is quite lengthy. Therefore, for the purposes of this thesis, instead of presenting several lines of straightforward, tedious, calculus, we choose to provide a Maple[1] worksheet which will verify that (A.7) and (A.9) are indeed correct.

The worksheet can be found at the end of this Appendix. In it, we begin by asking Maple to compute the integrals given in (A.6) and (A.8). The answers it provides are in a different form to those we give above, so we give Maple our solutions ((A.7) and (A.9)) and verify that the difference between Maple's answers and our's are zero.

Putting the results of (A.7) and (A.9) into (A.5) gives us the extra term for the price which we seek

$$\left(\frac{\alpha(M\nu^2)}{\sqrt{72\pi M}}\right)\left[(1 + \frac{(\log(K/S_0) - M\varepsilon)}{M\nu^2})K e^{-\frac{(\log(K/S_0)-M\varepsilon)^2}{2M\nu^2}} - S_0\sqrt{2\pi M\nu^2} e^{\frac{1}{2}M\nu^2 + M\varepsilon}\Phi(d_2)\right]. \quad \text{(A.10)}$$

This is the expression we give in Chapter 2, equation (2.10).

[1]Maple is an advanced mathematical software package, see http://www.maplesoft.com/

All that remains is to make our expression for α, (A.2), explicit. Observe that

$$\begin{aligned}
\alpha &= \mathbb{E}\left[\left(\frac{\zeta_1-\varepsilon}{\nu}\right)^3\middle|\zeta_1\in[a,b]\right]\\
&= \frac{1}{\nu^3}\mathbb{E}\left[\zeta_1^3-3\zeta_1^2\varepsilon+3\zeta_1\varepsilon^2-\varepsilon^3|\zeta_1\in[a,b]\right]\\
&= \frac{1}{\nu^3}\left[\mathbb{E}[\zeta_1^3|\zeta_1\in[a,b]]-3\mathbb{E}[\zeta_1^2|\zeta_1\in[a,b]]\varepsilon+2\varepsilon^3\right]. \qquad \text{(A.11)}
\end{aligned}$$

Except for $\mathbb{E}[\zeta_1^3|\zeta_1 \in [a,b]]$, we have explicit expressions for all the terms in (A.11). Notice that

$$\begin{aligned}
\mathbb{E}[\zeta_1^3|\zeta_1\in[a,b]] &= \int_{\mathbb{R}} x^3\mathbb{P}[\zeta_1^3\in dx|x\in\Gamma]\\
&= \int_{\mathbb{R}} x^3\frac{\mathbb{P}[\zeta_1\in dx]}{\mathbb{P}[\zeta_1\in[a,b]]}I_{\{x\in[a,b]\}}\\
&= \left(\frac{1}{\mathbb{P}[\zeta_1\in[a,b]]}\right)\left(\frac{1}{\sigma\sqrt{2\pi\delta}}\right)\int_a^b x^3 e^{\left(-\frac{(x-\mu\delta)^2}{2\sigma^2\delta}\right)}dx\\
&= \left(\frac{1}{\mathbb{P}[\zeta_1\in[a,b]]}\right)\left(\frac{1}{\sigma\sqrt{2\pi\delta}}\right)\left[\int_a^b x^2(x-\mu\delta)e^{\left(-\frac{(x-\mu\delta)^2}{2\sigma^2\delta}\right)}dx\right.\\
&\qquad\left.+\int_a^b x^2\mu\delta e^{\left(-\frac{(x-\mu\delta)^2}{2\sigma^2\delta}\right)}dx\right]\\
&= \left(\frac{1}{\mathbb{P}[\zeta_1\in[a,b]]}\right)\left(\frac{1}{\sigma\sqrt{2\pi\delta}}\right)\left[a^2\sigma^2\delta e^{-\frac{(a-\mu\delta)^2}{2\sigma^2\delta}}-b^2\sigma^2\delta e^{-\frac{(b-\mu\delta)^2}{2\sigma^2\delta}}\right]\\
&\qquad +2\sigma^2\delta\mathbb{E}[\zeta_1|\zeta_1\in[a,b]]+\mu\delta\mathbb{E}[\zeta_1^2|\zeta_1\in[a,b]]. \qquad \text{(A.12)}
\end{aligned}$$

Therefore (A.11) and (A.12) give,

$$\begin{aligned}
\alpha &= \left(\frac{\nu^{-3}}{\mathbb{P}[\zeta_1\in[a,b]]}\right)\left(\frac{1}{\sigma\sqrt{2\pi\delta}}\right)\left[\sigma^2\delta a^2\exp(-\frac{(a-\mu\delta)^2}{2\sigma^2\delta})-\sigma^2\delta b^2\exp(-\frac{(b-\mu\delta)^2}{2\sigma^2\delta})\right]\\
&\qquad+\left(\frac{1}{\nu^3}\right)\left[2\varepsilon(\varepsilon^2+\sigma^2\delta)+(\mu\delta-3\varepsilon)\mathbb{E}[(\zeta_1)^2|\zeta_1\in[a,b]]\right], \qquad \text{(A.13)}
\end{aligned}$$

which is the expression we give in Chapter 2, equation (2.9).

Maple Worksheet 1
This worksheet will verify that the integrals in (A.6) and (A.8) are indeed given by the expressions in (A.7) and (A.9).

```
> restart:
  assume(M>0,epsilon>0,K>S_0,S_0>0);
  normal_dist:=x->(1/sqrt(2*Pi))*int(exp(-t^2/2),t=-infinity..x);
```

First give maple the four integrals

```
> psi_1:=y->lambda*S_0*((y-M*epsilon)^2 /
  (M*gamma^2))*exp(y)*exp(-(y-M*epsilon)^2/(2*gamma^2*M));
  psi_2:=y->lambda*S_0*(y-M*epsilon)*exp(y)*exp(-(y-M*epsilon)^2/(2*
  gamma^2*M));
```

Next, ask Maple to evaluate the integrals (A.6) and (A.8)

```
> AA:=simplify(int(psi_1(y),y=-infinity..(log(K/S_0)))):
  BB:=simplify(int(psi_2(y),y=-infinity..(log(K/S_0)))):
```

We have supressed the ouput of the above equations. Unfortunately, Maple gives its answers in a form which looks very different to the expressions we have given in (A.7) and (A.9). So, to verify that they are indeed equal, we will now give Maple our expressions and then ask Maple to evaluate the difference between its answers (AA,BB) and our answers (A,B). Hopefully this should be zero.

```
> A:=simplify(K*lambda*(M*epsilon-log(K/S_0))*expand(exp(-expand((lo
  g(K/S_0)-M*epsilon)^2) /(2*gamma^2*M)))
  +BB+S_0*lambda*(int((exp(y)*exp(-( (y-M*epsilon)^2 /(2*gamma^2 *M)
  ) ) ),y=-infinity..(log(K/S_0) ) ) )):
  B:=-lambda*K*M*gamma^2*expand(exp(-expand((log(K/S_0)-M*epsilon)^2
  ) /(2*gamma^2*M)))+lambda*S_0*M*gamma^2*(2*Pi*M*gamma^2)^(1/2) *
  exp((1/2)*M*gamma^2 + M*epsilon)*normal_dist((log(K/S_0) -
  (M*gamma^2 + M*epsilon))/(gamma*sqrt(M))):
```

Now to evaluate the difference between the maple expression and our own.

```
> simplify(AA-A);
  simplify(BB-B);
```

Bibliography

Books

[1] T. Björk, *Arbitrage Theory in Continuous Time*, 1st Edition: Oxford University Press, Hardback ISBN 0198-7751-80 (1998)

[2] L. G. Douglas, *Yield Curve Analysis: the fundamentals of risk and return*, New York Institute of Finance, Hardback 0139-7245-67 (1988)

[3] J. A. Hartigan, *Clustering Algorithms*, 1st Edition:, Hardback ISBN 0471-3564-5X (1975)

[4] J. C. Hull, *Options, Futures, and Other Derivatives*, 3rd Edition: Prentice Hall International, Paperback ISBN 0132-6436-77 (1997)

[5] J. L. Jensen, *Saddlepoint Approximations*, 1st Edition: Oxford Clarendon Press, Hardback ISBN 0198-5229-59 (1995)

[6] J. E. Marsden and M. J. Hoffman, *Basic Complex Analysis*, 2nd Edition: Freeman, Hardback ISBN 0716-7181-46 (1987)

[7] M. Musiela and M. Rutkowski, *Martingale Methods in Financial Modelling*, 1st Edition: Springer, Hardback ISBN 3540-6147-7X (1998)

[8] S. N. Neftci, *An Introduction to the Mathematics of Financial Derivatives*, 1st Edition: Academic Press Limited, Hardback ISBN 0125-1539-02 (1996)

[9] V. V. Petrov, *Limit Theorems of Probability Theory*, 1st Edition: Clarendon Press, Hardback ISBN 0198-5349-9X (1995)

[10] W. H. Press, S. A. Teukolsky, W. T. Vettering and B. P. Flannery, *Numerical recipes in C: the art of scientific computing*, 2nd Edition: Cambridge University Press, Hardback ISBN 0521-4310-85 (1992)

[11] L. C. G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales, Vol 1: Foundations*, 2nd Edition: John Wiley & Sons, Hardback ISBN 0471-9506-10 (1994)

[12] D. Williams, *Probability with Martingales*, 1st Edition: Cambridge Mathematical Textbooks, Paperback ISBN 0521-4060-56 (1991)

Journal Articles and Papers

[13] F. Allen and D. Gale, *Financial Contagion*, Journal of Political Economy, **108** (2000), pp 1-33

[14] K. Amin and A. J. Morton, *Implied Volatility Functions in Arbitrage-Free Term Structure Models*, Journal of Financial Economics, **35** (1994), pp 141-180

[15] S. H. Babbs and N. Webber, *Term Structure Modelling Under Alternative Official Regimes*, Mathematics of Derivative Securities, Hardback ISBN 0521-5842-48 (1997), pp 394-422

[16] S. H. Babbs and N. J. Webber, *A Theory of the Term Structure with an Official Short Rate*, FORC working paper, **94/49** (1994)

[17] F. Black, E. Derman and W. Toy, *A One-factor Model of Interest Rates and its Application to Treasury Bond Options*, Financial Analysts Journal, **46** (1990), pp 33-39

[18] F. Black and M Scholes, *The Pricing of Options and Corporate Liabilities*, Journal of Political Economy, **81** (1973), pp 637-654

[19] M. I. Blejer, E. V. Feldman and A. Feltenstein, *Exogenous Shocks, Deposit Runs and Bank Soundness: A Macroeconomic Framework*, International Monetary Fund Working Paper, WP/97/91 (1997), pp 1-31

[20] P. P. Boyle and S. H. Lau, *Bumping Up Against the Barrier with the Binomial Method*, Journal of Derivatives (1994), 6-14

[21] A. Brace and M. Musiela, *A Multifactor Gauss Markov Implementation of Heath, Jarrow and Morton*, Mathematical Finance, **4** (1994), pp 259-283

[22] A. Brace, D. Gatarek and M. Musiela, *The Market Model of Interest Rate Dynamics*, Mathematical Finance, **7** (1995), pp 127-154

[23] D. T. Breeden, *Consumption, Production, Inflation and Interest Rates: A Synthesis*, Journal of Financial Economics, **16** (1986), pp 3-39

[24] M. J. Brennan and E. S. Schwartz, *A Continuous-Time Approach to the Pricing of Bonds*, Journal of Banking Finance, **3** (1979), pp 135-155

[25] M. J. Brennan and E. S. Schwartz, *An Equilibrium Model of Bond Pricing and a Test of Market Efficiency*, Journal of Financial Quantitative Analysis, **17** (1982), pp 301-329

[26] M. Broadie, P. Glasserman and S. Kou, *Connecting Discrete and Continuous Path-Dependent Options*, Finance and Stochastics, **3** (1999), pp 55-82

[27] M. Broadie, P. Glasserman and S. Kou, *A Continuity Correction for Discrete Barrier Options* , Mathematical Finance, **7** (1997), pp 325-349

[28] S. J. Brown and P. H. Dybvig, *The Empirical Implications of the Cox, Ingersoll, Ross Theory of the Term Structure of Interest Rates*, Journal of Finance, **XLI** (1986), pp 617-632

[29] R. H. Brown and S. M. Schaefer, *The Term Structure of Real Interest Rates and the Cox,Ingersoll and Ross Model*, Journal of Financial Economics, **35** (1994), pp 3-42

[30] W. Bühler, M Uhrig-Homburg, U. Walter and T. Weber, *An Empirical Comparison of Forward-Rate and Spot-Rate Models for Valuing Interest-Rate Options*, Journal of Finance, **LIV** (1999), pp 269-305

[31] P. Carr, *Two Extensions to Barrier Option Valuation*, Applied Mathematical Finance **2** (1995), pp 173-209

[32] K. C. Chan, G. A. Karolyli, F. A. Longstaff and A. B. Sanders, *An Empirical Comparison of Alternative Models of the Short-Term Interest Rate*, Journal of Finance **XLVII** (1992), pp 1209-1227

[33] R. Chen and L. Scott, *Maximum Likelihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates*, Journal of Fixed Income, (1993), pp 14-31

[34] G. M. Constantinides, *A Theory of the Nominal Term Structure of Interest Rates*, Review of Financial Studies, **5** (1992) pp 531-552

[35] J. C. Cox, C. Huang, *Optimal Consumption and Portfolio Policies when Asset Prices Follow a Diffusion Process*, Journal of Economic Theory, **49** (1989), pp 33-83

[36] J. C. Cox, J. E. Ingersoll and S. A. Ross, *An Intertemporal General Equilibrium Model of Asset Prices*, Econometrica, **53** (1985) pp 363-384

[37] J. C. Cox, J. E. Ingersoll and S. A. Ross, *A Theory of the Term Structure of Interest Rates*, Econometrica, **53** (1985) pp 385-407

[38] J. C. Cox, S. A. Ross & M. Rubinstein *Option pricing: a simplified approach*, Journal of Financial Economics, **7** (1979) pp 229-263

[39] H. E. Daniels, *Tail Probability Approximations*, International Statistical Review, **55** (1987), pp 37-48

[40] D. W. Diamond and P. H. Dybvig, *Bank Runs, Deposit Insurance and Liquidity*, Journal of Political Economy, **91** (1983), pp 401-419

[41] D. W. Diamond and R. G. Rajan, *Banks, Short Term Debt and Financial Crises: theory, policy, implications and applications*, Working paper University of Chicago and NBER, (2000)

[42] D. Duffie and R. Kan, *Multi-factor Interest Rate Models*, Philosophical Transactions Royal Society of London Series A, **347** (1994), pp 577-586

[43] B. Dumas, *Two-Person Dynamic Equilibrium in the Capital Market*, Review of Financial Studies, **2** (1989), pp 157-188

[44] P. H. Dybvig, *Bond and Bond Option Pricing Based on the Current Term Structure*, Working paper Washington University, (1988)

[45] M. Garman, *From Black Scholes to Black Holes*, Risk Magazine (book), ISBN 0951-6453-23 (1993), pp 89-92

[46] G. Gennotte and H. Leland, *Market Liquidity, Hedging and Crashes*, American Economic Review, **80** (1990), 999-1021

[47] H. Geman and M. Yor, *Pricing and Hedging Double-Barrier Options: A Probabilistic Approach*, Mathematical Finance, **6** (1996), pp 365-378

[48] M. R. Gibbons and K. Ramaswamy, *The Term Structure of Interest Rates: Empirical Evidence*, Review of Financial Studies, **6** (1993), pp 619-58

[49] M. B. Goldman, H. B. Sosin and M. A. Gatto, *Path Dependent Options: "Buy at the Low, Sell at the High"*, Journal of Finance, **34** (1979), pp 1111-1123

[50] J. M. Harrison and D. M. Kreps, *Martingales and Arbitrage in Multiperiod Securities Markets*, Journal of Economic Theory, **20** (1979), pp 381-408

[51] J. M. Harrison and S. R. Pliska, *Martingales and Stochastic Integrals in the Theory of Continuous Trading*, Stochastic Processes Applied, **11** (1981), pp 215-260

[52] D. C. Heath, R. A. Jarrow, & A. Morton, *Bond Pricing and the Term Structure of Interest Rates: A Discrete Time Approximation*, Journal of Financial Quantitative Analysis, **25** (1990), pp 419-440

[53] D. C. Heath, R. A. Jarrow, & A. Morton, *Contingent Claim Valuation with a Random Evolution of Interest Rates*, Review of Futures Markets, **9** (1990), pp 54-76

[54] D. Heath and D. Jara, *Term Structure Models based on Futures Prices*, preprint, (1998)

[55] J. P. Heritage and L. C. G. Rogers, *Large investors, takeovers and the rule of law*, To appear in Monte Carlo Methods and Applications, Springer Berlin, (2001).

[56] S. Heston, *Testing Continuous Time Models of the Term Structure of Interest Rates*, preprint (1989)

[57] T. S. Y. Ho and S. B. Lee, *Term Structure Movements and Pricing Interest Rate Contingent Claims*, Journal of Finance, **41** (1986), pp 1011-1029

[58] J. C. Hull and A. White, *Pricing Interest-Rate Derivative Securities*, Review of Financial Studies, **3** (1990), pp 573-592

[59] J. C. Hull and A. White, *One Factor Interest Rate Models and the Valuation of Interest Rate Derivative Securities*, Journal of Financial and Quantitative Analysis, **28** (1993), pp 235-254

[60] J. C. Hull and A. White, *A Note on the Models of Hull and White for Pricing Options on the Term Structure*, Journal of Fixed Income, (1995), pp 97-102

[61] J. C. Hull and A. White, *Using Hull and White Interest Rate Trees*, Journal of Derivatives, (1996), pp 26-36

[62] F. Jamshidian, *The One-Factor Gaussian Interest Rate Model: Theory and Implementation*, Working paper Merrill Lynch Capital Markets, (1988)

[63] F. Jamshidian, *Bond and Option Evaluation in the Gaussian Interest Rate Model*, Res Finance, **9** (1991), pp 131-170

[64] I. Karatzas, J. P. Lehoczky, J. P. Sethi and S. E. Shreve, *Explicit Solution of a General Consumption/Investment Problem*, Mathematics of Operations Research, **11** (1986), pp 261-294

[65] I. Karatzas, J. P. Lehoczky and S. E. Shreve, *Optimal Portfolio and Consuption Decisions for a Small Investor on a Finite Horizon*, SIAM Journal of Control and Optimization, **25** (1987), pp 1157-1186

[66] I. Karatzas, J. P. Lehoczky and S. E. Shreve, *Existence and uniqueness of multi-agent equilibrium in a stochastic, dynamic investment/consumption model*, Mathematics of Operations Research, **15** (1990), pp 80-128

[67] Kat and Verdonk, *Tree Surgery*, Risk Magazine, **8** (1995), pp 53-56

[68] N. Kunitomo and M. Ikeda, *Pricing Options with Curved Boundaries*, Mathematical Finance, **2** (1992), pp 275-298

[69] F. A. Longstaff, *A Nonlinear General Equilibrium Model of the Term Structure of Interest Rates*, Journal of Financial Economics, **23** (1989), pp 195-224

[70] F. A. Longstaff and E. S. Schwartz, *Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model*, Journal of Finance, **XLVII** (1992), pp 1259-1282

[71] R. E. Lucas, *Asset Prices in an Exchange Economy*, Econometrica, **46** (1978), pp 1429-1445

[72] P. Masson, *Contagion: macroeconomic models with equilibria*, Journal of International Money and Finance, **18** (1999), pp 587-602

[73] R. C. Merton, *Lifetime portfolio selection under uncertainty: the continuous-time case*, Review Economics and Statistics, **51** (1969), pp 247-257

[74] R.C. Merton, *Optimum Consumption and Portfolio Rules in a Continuous-time Model*, Journal of Economic Theory, **3** (1971), pp 373-413

[75] R. C. Merton, *An Intertemporal Capital Asset Pricing Model*, Econometrica, **41** (1973), pp 867-880

[76] R. C. Merton, *Theory of rational option pricing*, Bell Journal of Economics and Management Science, **4** (1973), pp 141-183

[77] F. S. Mishkin, *Lessons from the Asian crisis*, Journal of International Money and Finance, **18** (1999), pp 709-723

[78] J. M. Moraleda and A. Pelsser, *Forward versus Spot Interest Rate Models of the Term Structure: An Empirical Comparison*, Journal of Derivatives, (2000), pp 9-21

[79] N. D. Pearson and T. Sun, *Exploiting the Conditional Density in Estimating the Term Structure: An Application to the Cox, Ingersoll, and Ross Model*, Journal of Finance, **XLIX** (1994), pp 1279-1304

[80] D. Rich, *The Mathematical Foundations of Barrier Option Pricing Theory*, Advances in Futures and Options Research, **7** (1994), pp 267-312

[81] L. C. G. Rogers and E. J. Stapleton, *Fast Binomial Pricing*, Finance and Stochastics, **2** (1998), pp 3-17

[82] L. C. G. Rogers and O. Zane, *Saddle-point Approximations to Option Pricing*, preprint (1997) University of Bath, England

[83] L. C. G. Rogers, *Which Model for the Term-Structure of Interest Rates Should One Use?*, Mathematical Finance, **65** (1995), pp 93-116

[84] L. C. G. Rogers, *The Potential Approach to the Term-Structure of Interest Rates and Foreign Exchange Rates*, Mathematical Finance, **7** (1997), pp 157-176

[85] L. C. G. Rogers and O. Zane, *Fitting Potential Models to Interest Rate and Foreign Exchange Data*, Vasicek and Beyond (book), ISBN 1899-3325-02 (1996), pp 327-342

[86] L. C. G. Rogers and F. A. Yousaf, *Markov Chains and the Potential Approach to Modelling Interest Rates and Exchange Rates*, To appear in the Volume of the Bachelier Proceedings 2000-2001

[87] L. C. G. Rogers and F. A. Yousaf, *Examining the Effects of Shocks to a Market Equilibrium*, working paper, University of Bath (2001)

[88] M. Rubinstein and E. Reiner, *Breaking Down the Barriers*, Risk Magazine, **4** (1994), pp 28-35

[89] J. Saa-Requejo, *The Dynamics and the Term Structure of Risk Premia in Foreign Exchange Markets*, preprint (1993)

[90] O. Vasicek, *An Equilibrium Characterisation of the Term Structure*, Journal of Financial Economics, **5** (1977), pp 177-188

[91] J. Wang, *The Term Structure of Interest Rates in a Pure Exchange Economy with Heterogeneous Investors*, Journal of Financial Economics, **41** (1996), pp 75-110

[92] F. A. Yousaf, *Pricing a Knockdown Option*, submitted for publication to the Journal of Applied Probability (March 2001)